

KARADENİZ TECHNICAL UNIVERSITY * INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE

MASTER'S PROGRAM IN APPLIED LINGUISTICS

**VOICES FROM EFL TEACHERS: INTEGRATING LISTENING AND SPEAKING
COMPONENTS IN ASSESSMENT**

MASTER'S THESIS

Rabia ÇAPKIN ACAR

OCTOBER-2025

TRABZON

KARADENİZ TECHNICAL UNIVERSITY * INSTITUTE OF SOCIAL SCIENCES

**DEPARTMENT OF WESTERN LANGUAGES AND LITERATURE
MASTER'S PROGRAM IN APPLIED LINGUISTICS**

**VOICES FROM EFL TEACHERS: INTEGRATING LISTENING AND SPEAKING
COMPONENTS IN ASSESSMENT**

MASTER'S THESIS

Rabia ÇAPKIN ACAR

ORCID:

Supervisor: Assistant Prof. Dr. Hasan SAĞLAM

OCTOBER-2025

TRABZON

APPROVAL

Upon the submission of the dissertation, **Rabia ÇAPKIN ACAR** has defended the study **“Voices from EFL Teachers: Integrating Listening and Speaking Components in Assessment”** in partial fulfillment of the requirements for the degree of Master of Arts in Applied Linguistics at Karadeniz Technical University, and the study has been found fully adequate in scope and quality as a thesis **unanimously on 14/10/2025**.

Committee Member	Decision			Signature
Title - Name and Surname	Mission	Accept	Refuse	
Prof. Dr. Mustafa Naci KAYAOĞLU	Chair	☐	☐	
Prof. Dr. Turgay HAN	Member	☐	☐	
Asst. Prof. Dr. Hasan SAĞLAM	Member	☐	☐	

Approval of the Graduate School of Social Sciences

Prof. Dr. Haydar AKYAZI
Director

DECLARATION OF ORIGINALITY

- **I, Rabia ÇAPKIN ACAR**, hereby confirm and certify that;
- I am the sole author of this work, and I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution,
- this work contains no material that has been submitted or accepted for a degree or diploma in any other university or institution,
- all data and findings in the work have not been falsified or embellished,
- this is a true copy of the work approved by my advisor and thesis committee at Karadeniz Technical University, including final revisions required by them,
- I understand that my work may be electronically checked for plagiarism by the use of plagiarism detection software and stored on a third party's server for eventual future comparison,
- I take full responsibility in case of non-compliance with the above statements in respect of this work.

Signature:

Date: 14.10.2025

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to everyone who believed in me, helped, and encouraged me throughout the process of writing this thesis. First and foremost, I am deeply grateful to my thesis advisor, Asst. Prof. Dr. Hasan SAĞLAMER, whose knowledge, patience, experience, and continuous guidance have been invaluable throughout this journey. I also owe a debt of gratitude to Prof. Dr. Mustafa Naci KAYAOĞLU for his fatherly support, devoted feedback, and insightful contributions, which greatly helped me refine and strengthen my thesis. My sincere thanks also go to Prof. Dr. Turgay HAN for his valuable feedback, which significantly contributed to the development of my study.

Furthermore, I would like to express my sincere appreciation to all my colleagues at the secondary schools in Ortahisar, Trabzon, who generously shared their valuable insights and experiences, enriching this research in meaningful ways.

I am profoundly grateful to my parents, Gülüşan ÇAPKIN and Mehmet Salih ÇAPKIN, for their unconditional love, patience, and encouragement throughout my academic journey.

Lastly, I would like to extend my deepest thanks to my husband, Doğan ACAR, for his endless support and understanding whenever I needed it, and to my son, Asaf Uygur, whose joy and love have been my greatest source of motivation during this process.

October, 2025

Rabia ÇAPKIN ACAR

TABLE OF CONTENTS

ACKNOWLEDGEMENT	IV
TABLE OF CONTENTS	V
ÖZET	VII
ABSTRACT.....	VIII
LIST OF TABLES.....	IX
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XI
INTRODUCTION.....	1-5

CHAPTER ONE

1. LITERATURE REVIEW.....	6-25
1.1. Theoretical Background.....	6
1.1.1. Discrete Point Testing to Communicative Testing	8
1.1.2. Introduction of Performance-Based Assessment in Speaking and Listening	9
1.1.3. Integrated Skills Testing.....	9
1.2. Assessment	11
1.3. Types of Assessment.....	11
1.3.1. Summative and Formative Assessment	11
1.3.2. Informal and Formative Assessment	12
1.3.3. Classroom Assessment	12
1.4. Assessment Literacy and Assessment Literacy in Türkiye.....	13
1.5. Listening	16
1.5.1. Listening Assessment	19
1.6. Speaking	20
1.6.1. Speaking Assessment	23

CHAPTER TWO

2. METHODOLOGY.....	26-38
2.1. Research Design (Case study/mixed method)	26
2.2. Quantitative Research	28
2.2.1. Participants and Setting	28
2.3. Qualitative Research	30
2.3.1. Participants and Setting	30
2.4. Data Collection Instruments.....	31
2.4.1. Questionnaire	31
2.4.2. Semi-Structured Interview.....	32
2.5. Data Analysis.....	33
2.5.1. Quantitative Data Analysis.....	34
2.5.2. Qualitative Data Analysis.....	34
2.6. Triangulation.....	34
2.7. Ethical Considerations	35
2.8. Validity and Reliability	36
2.9. Positionality in the Study: Being an Insider as a Researcher.....	37

CHAPTER THREE

3. FINDINGS.....	39-85
3.1. Quantitative Data	39
3.2. Qualitative Data	78

CHAPTER FOUR

4. DISCUSSIONS.....	86-92
4.1. Quantitative Data	86
4.2. Qualitative Data	90

CONCLUSION AND IMPLICATIONS.....	93
---	-----------

REFERENCES.....	98
------------------------	-----------

APPENDICES	113
-------------------------	------------

CIRRICULUM VITAE.....	129
------------------------------	------------

ÖZET

Bu çalışma, 2023-2024 akademik yılından itibaren Türkiye’de ortaokul düzeyinde uygulanmaya başlanan dinleme ve konuşma becerilerinin değerlendirilmesine ilişkin İngilizce öğretmenlerinin görüşlerini incelemektedir. Dil öğretimi geleneksel olarak okuma ve yazma becerilerine ağırlık vermiş olsa da dinleme ve konuşmanın değerlendirilmesi son dönemde yeni bir odak alanı hâline gelmiştir. Çalışmanın amacı, öğretmenlerin bu değerlendirmelere ne kadar hazır olduklarını, kullandıkları yöntemleri, karşılaştıkları zorlukları ve geliştirmeye yönelik önerilerini ortaya koymaktır. Araştırma, karma yöntem yaklaşımını benimsemekte; nicel veriler anket yoluyla, nitel veriler ise yarı yapılandırılmış görüşmeler aracılığıyla toplanmıştır. Katılımcılar, Trabzon İlinin Birinci Hizmet Bölgesi’nde görev yapan Ortaokul İngilizce öğretmenleridir. Bu çalışma, öğretmenlerin dinleme ve konuşma becerilerinin değerlendirilmesinde kullandıkları teknikler, bu tekniklerin geçerlilik ve güvenilirliğine ilişkin algıları ve mevcut uygulamalara yönelik önerileri hakkında önemli bulgular sunmaktadır. Araştırma, Türk ortaokullarında görev yapan İngilizce öğretmenlerinin bakış açısından hareketle, dinleme ve konuşma becerilerinin değerlendirilmesinin entegrasyonunu ve bunun Türkiye bağlamındaki etkilerini inceleyen güncel ve önemli bir konuyu ele alması bakımından dikkate değerdir. Ayrıca, çalışma aracılığıyla, İngilizce öğretmenlerinin geleneksel ölçme-değerlendirme yöntemlerinden iletişimsel ve performansa dayalı ölçme sürecine geçişe yönelik hazır bulunuşluk düzeylerinin de incelenmesi amaçlanmaktadır. Bulgular, öğretmenlerin özellikle ölçme araçlarının geliştirilmesi, standartlaştırma, değerlendirme güvenilirliği ve zaman-mekân sınırlamaları gibi konularda çeşitli zorluklarla karşılaştıklarını göstermektedir. Bununla birlikte öğretmenler, bu becerilerin ölçülmesinin öğrenci motivasyonunu artırdığını, iletişimsel yetilerini geliştirdiğini ve öğretim etkinliğini olumlu yönde etkilediğini belirtmiştir. Elde edilen bulgular, bu alanda öğretmenlere yönelik hizmet içi eğitimlerin ve destek mekanizmalarının artırılması gerekliliğini vurgulamaktadır. Sonuç olarak, bu araştırma, dinleme ve konuşma becerilerinin değerlendirilmesinin önemini ve mevcut durumunu ortaya koyarak Türkiye’de yabancı dil öğretimi ve değerlendirmesi alanına hem politika yapıcılar hem de eğitimciler için anlamlı bir katkı sağlamayı hedeflemektedir.

Anahtar Kelimeler: Dinleme ve Konuşma Becerileri, İngilizce Öğretmenleri Görüşleri, Ölçme ve Değerlendirme, Performans Temelli Değerlendirme, Ölçme Okuryazarlığı

ABSTRACT

This study explores English teachers' perspectives on assessing listening and speaking skills, which has been implemented at the secondary school level in Türkiye since the 2023-2024 academic year. Although language teaching traditionally emphasizes reading and writing skills, the evaluation of listening and speaking has become a new focus area. The study aims to assess teachers' readiness for these assessments, the methods they employ, the challenges they encounter, and their suggestions for improvement. It uses a mixed-methods approach, collecting quantitative data through questionnaires and qualitative data through semi-structured interviews. Participants include Secondary School English Teachers from the First Service Area in Trabzon Province. The study provides insights into the assessment techniques teachers use for listening and speaking, their perceptions of the validity and reliability of these methods, and their recommendations for current practices. This study is significant as it attempts to investigate a timely issue, the integration of speaking and listening assessment, and its effects in the Turkish context, taking the perspectives of Turkish Secondary School EFL teachers as a basis. It aims to shed light on a gap in the Turkish curriculum, portraying real-life challenges and experiences of Secondary School EFL teachers in the Turkish context on the newly implemented assessment of speaking and listening skills. Through this study, it is also aimed to examine the readiness of EFL teachers for the transition from traditional assessment methods to communicative and performance-based assessment processes. Results indicate that teachers face several challenges, especially regarding the development of assessment tools, standardization, assessment reliability, and time and space limitations. Teachers also report that measuring these skills can boost student motivation, enhance communicative abilities, and positively influence teaching effectiveness. The findings highlight the need for increased in-service training and support for teachers in this area. Ultimately, this research aims to contribute to the field of foreign language teaching and assessment in Türkiye by highlighting the importance and current state of listening and speaking skill assessment for policymakers and educators alike.

Keywords: Listening and Speaking Skills, EFL Teachers' Perspectives, Assessment and Evaluation, Performance-Based Assessment, Assessment Literacy

LIST OF TABLES

Table No	Table Name	Page Nr.
1	Demographic Information about Participants	29
2	The Cronbach's Alpha of the Questionnaire, Listening and Speaking Assessment	37
3	Statistics and Frequency Values for Q-3 & Q-37	42
4	Statistics and Frequency Values for Q-4 & Q-38	44
5	Statistics and Frequency Values for Q-5 & Q-39	45
6	Statistics and Frequency Values for Q-7 & Q-41	46
7	Statistics and Frequency Values for Q-8 & Q-42	47
8	Statistics and Frequency Values for Q-9 & Q-43	49
9	Statistics and Frequency Values for Q-10 & Q-44	50
10	Statistics and Frequency Values for Q-11 & Q-45	51
11	Statistics and Frequency Values for Q-13, Q-14, Q-47 & Q-48	54
12	Statistics and Frequency Values for Q-19 & Q-53	58
13	Statistics and Frequency Values for Q-24 & Q-58	64
14	Statistics and Frequency Values for Q-28 & Q-62	68
15	Statistics and Frequency Values for Q-29 & Q-63	70
16	Statistics and Frequency Values for Q-33 & Q-67	75
17	Statistics and Frequency Values for Q-34 & Q-68	77
18	Thematic Coding Table	78
19	Codes and Quotes for Theme 1: Contributions to Students	79
20	Codes and Quotes for Theme 2: Challenges and Disadvantages Theme	81
21	Codes and Quotes for Theme 3: Teachers' Observations Theme	81
22	Codes and Quotes for Theme 4: Educational Contributions Theme	82
23	Codes and Quotes Theme 5: Views on Assessment Methods Theme	84
24	Codes and Quotes for Theme 6: General Attitudes & Suggestions Theme	84

LIST OF FIGURES

Figure No.	Table Name	Page Nr.
1	Statistics and Frequency Values for Q-1 and Q-35	40
2	Statistics and Percentage Values for Q-2 & Q-36	41
3	Statistics and Frequency Values for Q-12 & Q-46.....	52
4	Statistics and Frequency Values for Q-15 & Q-49.....	55
5	Statistics and Frequency Values for Q-17 & Q-51.....	57
6	Statistics and Frequency Values for Q-18 & Q-52.....	57
7	Statistics and Frequency Values for Q-20 & Q-54.....	60
8	Statistics and Frequency Values for Q-21 & Q-55.....	61
9	Statistics and Frequency Values for Q-22 & Q-56.....	62
10	Statistics and Frequency Values for Q-23 & Q-57.....	64
11	Statistics and Frequency Values for Q-25 & Q-59.....	66
12	Statistics and Frequency Values for Q-27 & Q-61	67
13	Statistics and Frequency Values for Q-30 & Q-64.....	72
14	Statistics and Frequency Values for Q-31 & Q-65.....	73
15	Statistics and Frequency Values for Q-32 & Q-66.....	74

LIST OF ABBREVIATIONS

CALI	: Classroom Assessment Literacy Inventory
CBA	: Classroom-Based Assessment
CLT	: Communicative Language Teaching
EFL	: English as a Foreign Language
ELT	: English Language Teaching
ELTE	: English Language Testing and Evaluation
LMS	: Learning Management Systems
METU	: Middle East Technical University
PBA	: Performance-Based Assessment
TBLA	: Task-Based Language Assessment
TOEFL	: Test of English as a Foreign Language
TOEFL-IBT	: Test of English as a Foreign Language – Internet-Based Test

INTRODUCTION

In recent years, the assessment of listening and speaking skills has gained increasing attention in the field of English language education, reflecting a broader shift toward communicative and performance-based approaches to language learning. However, in many educational contexts, including Türkiye, the evaluation of these oral skills has often been overshadowed by a stronger focus on reading and writing. Recognizing this imbalance, the Turkish Ministry of National Education introduced formal listening and speaking assessments into the secondary school curriculum during the 2023–2024 academic year. This reform represents a significant step toward aligning national assessment practices with the principles of communicative teaching. In this context, the present study investigates how English teachers in Trabzon perceive, interpret, and implement this new policy, aiming to uncover their experiences, challenges, and perceptions regarding the assessment of listening and speaking skills within the evolving landscape of English language education in Türkiye.

Background and Rationale of the Study

Listening and speaking are essential aspects of communication (Sabina, 2018). These skills are necessary for real-world communication, yet their assessment has not received as much attention as reading and writing (Yu et al., 2021). The assessment of listening and speaking skills is often perceived as challenging because it is based on real-life performance and factors that influence it, such as context, environment, and proficiency levels. Following the introduction of formal assessments for listening and speaking skills in secondary schools during the 2023–2024 academic year, it became necessary to develop an evidence-based understanding of how these skills are assessed and implemented by teachers (Han and Elçiçek, 2021; *Milli Eğitim Bakanlığı Ölçme ve Değerlendirme Yönetmeliği*, 2023). This recent change in the education system has brought several challenges for teachers. One of the most prominent issues is that many teachers are not adequately prepared to assess listening and speaking skills, as they have not received sufficient training in relevant assessment techniques (Maryslessor et al., 2014). It is readily apparent that listening and speaking are more complex abilities compared to other skills, and students tend to have more difficulty with these skills, as they necessitate the real-life production of language. Although there are studies on the assessment of listening and speaking skills (Caruso et al., 2017; Flowerdew and Miller, 2012; Sarwar et al., 2014; Srikaew et al., 2015; Zaim and Arsyad, 2020), there is limited research on the difficulties teachers face when making assessments in real classroom settings. Another factor contributing to the challenges is the consideration that teachers

may not be adequately equipped with the necessary techniques and methods for assessment (Oo et al., 2023; Rajput et al., 2020). Their point of view, experiences, and challenges in real classroom assessment periods need to be explored systematically to provide effective solutions.

The rationale of this study is to investigate the teachers' perceptions of the reliability and validity of their assessments of these skills and to identify potential ways to enhance their assessment processes, thereby raising awareness of their shortcomings through this study and the literature reviewed. This study is significant because it may contribute to a timely issue of integrating speaking and listening skills in secondary schools, focusing on the voices of EFL teachers regarding the validity and reliability of their assessment practices, and aiming to reveal the challenges and needs arising from this new integration. The study also examines methods for identifying potential opportunities for the professional development of teachers and assessment procedures in secondary schools that enhance these skills. By doing so, this study may inform policymakers' upcoming efforts to enhance teachers' readiness for communicative and performance-based assessments of speaking and listening. Such inquiry is particularly significant for language assessment in Turkish secondary schools, where speaking and listening often receive less systematic attention (Çelik and Sur, 2025). By assessing teachers' perspectives through both interviews and survey data, and interpreting these findings in light of the relevant literature, it becomes possible to gain a deeper understanding of the existing challenges and opportunities. Meeting this need has the potential to inform policy decisions and classroom practices alike, ultimately enhancing the quality of English language assessment and instruction in Türkiye and comparable educational contexts.

Statement of the Problem

In recent years, speaking and listening assessments have gained significant importance; many studies, in particular, emphasize the need for communicative and performance-based evaluations. Contrary to traditional language-assessment techniques, which place more emphasis on learning vocabulary and grammar rules, communicative language assessment firmly acknowledges that the main goal of language learning is to enable smooth communication with others (Green, 2020). Therefore, communicative language assessment emphasizes the importance of enabling students to use language in real-world contexts. The current pedagogical approach of task-based language teaching, which emphasizes the application of language abilities in real-world, practical contexts, is also aligned with communicative language evaluation.

All over the world, people have been seeking authentic assessment techniques that accurately reflect real-life performance in listening and speaking skills. However, their assessment practice appears to have been neglected in comparison to other language skills, as reviewed in the literature, even though listening and speaking abilities are crucial in language learning (Qayoom et al., 2021).

Although the importance of these skills is well established, research shows that assessing them, especially speaking, which involves real-life performance and multiple influencing factors—remains a challenge. Beenaf (2016) emphasizes that many teachers recognize their need for training in assessing these skills and express a strong desire to receive such training.

In Türkiye, speaking and listening skills were officially integrated into the national curriculum in the 2023–2024 academic year. The assessment of these skills also began at the secondary school level during the same period, representing a new practice for many teachers (Akdoğan & Toprak, 2024; Çelik & Sur, 2025). Although these skills had previously been included in classroom activities to some extent, they were not formally graded by teachers, or when they were, the grading policies and weightings lacked clarity. These formal assessments for listening and speaking in secondary schools bring about a significant change in language education policy. However, there is insufficient research and established procedures on how teachers should conduct these exams in practice, particularly in written formats. Regarding the fact that teachers may not be professionally ready to assess the listening and speaking skills of their students in a formal way, this study aims to examine whether teachers are aware of and ready for assessment techniques for listening and speaking skills. For this purpose, the study seeks to portray teachers' experiences, difficulties, and perceptions regarding the assessment of these skills.

Although considerable research exists on the assessment of these skills, there is a notable lack of research on the real-life application of these skills and their evaluation in the Turkish context. Therefore, it is essential to explore the challenges and realities of assessing these skills in Türkiye's secondary schools, with a focus on teachers' experiences, needs, and recommendations to enhance the effectiveness of speaking and listening assessment.

Purpose of the Study

The purpose of this study is to explore EFL teachers' perceptions regarding the integration of listening and speaking assessment as a formal component of the Turkish secondary school curriculum. Since education authorities have recently incorporated the assessment of these skills, it is necessary to examine how teachers perceive and implement them in classroom settings. The study aims to evaluate the effectiveness and practicality of this newly adopted assessment approach in measuring students' real-life competencies in listening and speaking.

By investigating teachers' perspectives on the reliability and validity of current practices, this study seeks to bridge the gap between research-based knowledge and classroom realities in the assessment of oral skills. Assessing speaking, in particular, often differs from authentic communication contexts and theoretical understandings, which highlights the importance of exploring teachers' readiness to assess these skills effectively. With the inclusion of listening and

speaking in formal assessments, understanding teachers' preparedness becomes crucial for identifying challenges and developing appropriate solutions that connect pedagogical theory with classroom practice.

Drawing on both quantitative and qualitative data, as well as insights from the existing literature, the study aims to provide comprehensive and reliable findings about the experiences, perceptions, and challenges of EFL teachers in assessing students' listening and speaking skills in Turkish secondary schools. It also investigates teachers' levels of readiness and their awareness of suitable assessment techniques for evaluating these skills. By addressing these aspects, the study aspires to close the gap between the challenges faced in practice and the solutions proposed in research, thereby contributing to the development of more effective, valid, and contextually relevant assessment processes grounded in teachers' experiences and perspectives.

Research Questions

Based on the study's aims, the following research questions were determined. The research question is divided into two parts: the main research question is followed by minor questions.

Major Question:

1. What assessment practices do EFL teachers use to evaluate students' listening and speaking skills?
2. What are EFL teachers' perspectives on the integration of listening and speaking skills into secondary school assessment processes?

Significance of the Study

This study is significant for several reasons. First, it fills a crucial gap in the literature—addressing a notable void concerning the formal assessment of listening and speaking skills in secondary schools in Türkiye. This relatively new area requires further investigation. Although there is a consensus on the importance of these skills for oral communication, limited research has explored how they are or should be formally assessed, particularly within the Turkish educational context. Because these abilities are essential for successful interaction, understanding both current and ideal assessment practices can offer valuable insights into improving language assessment approaches for listening and speaking skills, especially in Turkish secondary schools.

Moreover, this study may provide EFL teachers with opportunities to become familiar with innovative methods of assessing students. The findings can also enable them to learn from their colleagues' experiences and perspectives, helping them develop more effective assessment practices. Through the literature review presented in this study, teachers may also gain awareness

of potential limitations or misconceptions in their current assessment techniques. In addition, the results can help establish more effective and practical assessment procedures that meet students' communicative needs. By incorporating teachers' opinions, this study seeks to contribute to the development of valid and reliable assessment instruments, thereby enhancing the overall quality of language teaching.

Delimitation of the Study

This study is delimited to a specific educational context, population, and scope that align with its research aims. The research focuses on secondary school English teachers working in the first service area of Ortahisar, Trabzon, during the 2023–2024 academic year. The selection of this group was intentional, as these teachers directly experienced the newly introduced policy mandating the assessment of listening and speaking skills at the secondary school level in Türkiye. Therefore, the study does not include teachers from primary or high schools, nor does it investigate the perspectives of students, administrators, or policymakers.

The study is further delimited to the assessment of listening and speaking skills, excluding other language skills such as reading and writing. This focus was chosen to address the gap in research concerning oral skill assessment, as listening and speaking evaluation had only recently been implemented in the Turkish EFL curriculum. Additionally, the study employs a mixed-methods case study design, utilizing questionnaires and semi-structured interviews as the primary data collection instruments. Other qualitative methods, such as classroom observation, document analysis, or focus group interviews, were deliberately excluded to maintain feasibility within the study's time frame.

Moreover, the geographical scope of the study is confined to Trabzon-Ortahisar, a region that offers accessibility and represents a realistic local context for implementing the new curriculum reform. The study period spans the 2023–2024 school year, during which the policy was first implemented, to capture teachers' initial perceptions and experiences. Therefore, the findings reflect the conditions and practices specific to this time and region. By defining these boundaries, the study aims to ensure a focused, manageable, and contextually meaningful investigation of EFL teachers' perspectives on integrating listening and speaking assessments into secondary education in Türkiye.

CHAPTER ONE

1. LITERATURE REVIEW

The concept of language assessment has undergone a significant transformation in tandem with changes in English language teaching. While grammar and vocabulary were the primary focus of language teaching, communicative approaches have gained prominence since the 1970s. This led students to apply language in real-life situations and to integrate their communication skills into the assessment process. Therefore, there has been a shift from discrete-point tests to communicative, performance-based assessment methods that consider multiple skills simultaneously.

Along with these developments, the concept of assessment literacy has also gained importance. Teachers are expected to plan assessment processes effectively, select appropriate tools, and use the results to inform and improve their teaching. Assessment is now seen not only as a tool for grading students, but also as a process that supports learning. This chapter addresses the historical development of language assessment, the different types of assessment used, and the importance of teachers' assessment literacy, with a particular focus on the context of English language teaching in Türkiye.

1.1. Theoretical Background

Reading and writing were the main focuses of teaching English as a foreign language in the past, but following the 1970s, listening and speaking became critical for reading and writing, as the teaching population recognized the importance of communicative teaching. Speaking has been an essential part of language proficiency development since the beginning of communicative instruction in the 1970s. Listening, like speaking, regained prominence in the language classroom after being overlooked for a long time. Listening and speaking skills are not only important but also challenging to assess as the researchers suggest (Fakhrurriana and Herdina, 2024; Lankara et al., 2024; Palmer, 2014; Richards, 2008; R. B. Rubin, 2009; Slamet and Mukminatien, 2024; Turner, 1998); they could not happen suddenly; they require interactive processes, which many of the teachers lack in assessing these skills. In recent years, there has been a shift away from traditional standardized exams in favor of innovative methods that assess learners' development (Awad and Al Adwan, 2024; Bashir, 2024).

As language instruction has grown more student-centered, other assessment methods have gained popularity. Over two decades ago, Darling-Hammond et al. (1995) proposed the use of

alternative assessment methodologies. They provided "performance assessments" instead of traditional assessment procedures (Darling-Hammond et al., p. 15). Researchers later categorized assessments. Whereas Reeves (2000) divides alternative assessment methods into "performance or authentic assessment and portfolio assessment," Gilligan (2007) adapts Erickson's (2001) categorization as standard-based assessments, performance assessments, portfolio assessments, and authentic assessments. Performance assessments involve, for example, role-plays, dialogues, and presentations that focus on students' actual performance rather than linguistic competence, which is much more student-centered and acceptable to today's researchers. According to Stiggins (2002), teachers first need to understand the full range of assessment types and their applications. Users of assessment—such as teachers, students, school administrators, policymakers, education authorities, and parents—exist at various levels, from classrooms to educational institutions and administrative bodies. These users require different types and timings of information about student achievement to fulfill their respective roles. There is no single assessment model that fits all purposes. Assessment-literate teachers recognize and appreciate these differences. Second, teachers must clearly identify the learning goals their students are expected to achieve. Different kinds of achievement—such as mastery of subject knowledge, reasoning ability, performance skills, or product creation—require different assessment methods. Assessment professionals understand that the first question to ask is, "What do I want to assess?" Only after answering this question, they decide on the most appropriate evaluation method. Third, teachers need to be prepared to measure student performance using a variety of assessment formats, including selected response questions, essays, performance-based tasks, and personal communication assessments. They must be able to choose the most suitable method for their specific context and understand when and why each is appropriate. Teachers today are expected to be far more assessment-aware than they were in the past. Finally, teachers must learn how to design assessment activities that quickly and accurately measure student performance using the chosen approach. Assessment-literate individuals collect sufficient and relevant evidence of student achievement to make confident and valid judgments.

The study suggests that teachers need to be aware of potential sources of bias that may influence any assessment and to take reliable measures to prevent misrepresentation of assessment results. Moreover, assessment-literate teachers recognize the connection between assessment and student motivation. They understand how to involve students in assessment processes and communicate effectively, thereby transforming these into trust-building instructional practices. They also recognize that meaningful student involvement substantially increases the likelihood of academic success.

Considering all these factors, secondary school teachers in Türkiye face significant challenges due to their limited experience and lack of training in the required assessment techniques. Many teachers found themselves unexpectedly involved in this process after the introduction of speaking and listening assessments in secondary schools during the 2023–2024

academic year. Therefore, this study aims to bridge the existing gap between research findings, theoretical knowledge, and classroom practices by employing both qualitative and quantitative data to explore teachers' perspectives and provide a clear vision for implementing reliable and valid assessment practices.

Reading and writing were the primary focuses of teaching English as a foreign language in the past, but following the 1970s, listening and speaking became critical for reading and writing as the importance of communicative teaching was recognized by the teaching community (Farooq, 2015). Speaking has been an essential component of language proficiency development since the introduction of communicative instruction in the 1970s. Listening, like saying, regained prominence in the language classroom after being overlooked for an extended period. Listening plays a significant part in language instruction by providing learners with valuable input (Mart, 2020). Listening and speaking skills are not only essential but also challenging to assess, as researchers suggest; they cannot happen suddenly; they require interactive processes, which many teachers lack in assessing these skills. In recent years, there has been a shift away from traditional standardized exams in favor of creative ways that assess learners' development.

1.1.1. Discrete Point Testing to Communicative Testing

The shift from discrete point tests to communicative tests presents a significant, dramatic paradigmatic shift in language assessment approaches. This shift is driven by the need for more inclusive assessment of learners' ability to use the language in real-life situations. Discrete point testing refers to an approach that assesses individual components of language—such as grammar, vocabulary, or pronunciation—separately, assuming that the sum of these parts can represent overall proficiency. That is why discrete point tests fail to reflect the communicative ability of learners. Therefore, this approach, which was common in the 20th century, has been replaced over time by integrative and communicative methods (Farhady, 2018; Yuzar, 2020).

Communicative Testing, in contrast, aims to evaluate how effectively learners can use language to communicate meaning in authentic contexts, focusing on fluency, appropriateness, and interaction rather than isolated linguistic knowledge. Communicative assessment focuses on the functional use of language, ensuring compatibility with current teaching approaches that emphasize learners' communicative ability and comfort (Yuzar, 2020). According to Salekhova et al. (2019), three main trends currently influence multilingual proficiency assessments: native language awareness, differential scoring based on language background, and translanguaging. This reveals the inadequacy of monolingual assessment models (Hemsley et al., 2014).

1.1.2. Introduction of Performance-Based Assessment in Speaking and Listening

In recent years, performance-based assessment (PBA) has gained increasing importance in evaluating speaking and listening skills. This increasing interest is linked to the realization that learners' ability to use language effectively in real-life contexts needs to be measured. While traditional assessments often focus on rote knowledge or discrete language elements, the PBA approach aims to measure learners' language proficiency through real-life tasks (Noroozi and Taheri, 2021). Task-Based Language Assessment (TBLA), which is at the center of this transformation, extends beyond language knowledge to reveal how effectively learners utilize language in task-based interactions (Noroozi and Taheri, 2021).

One of the key advantages of PBA is that it supports the collaboration and active participation of learners. Research indicates that strategies such as explicit guidance, cooperative learning, and gradual support have a positive impact on language development (Lunar et al., 2023). Task-based activities enable learners to collaborate and receive feedback from one another. This process supports the development of both speaking and listening skills (Lunar et al., 2023). In this way, the learning process is socialized, and language skills are reinforced in a more meaningful context. Effective implementation of PBA requires teachers to have assessment literacy. Especially in English as a foreign language (EFL), teachers need specialized training in terms of both subject matter knowledge and pedagogical strategies (Firoozi et al., 2019). A practical assessment should include not only measurement but also the process of giving feedback. In this way, teachers can adapt their teaching strategies according to students' performance and control the learning process (Prastikawati et al., 2022). Technology integration is also an important parameter supporting PBA. Digital tools can be used to create interactive environments that increase students' participation and allow students to do tasks (Nurharjanto and Widyanoro, 2020). In addition, the use of such technologies increases students' sense of self-efficacy and motivation, making them more successful, especially in high-anxiety tasks with speaking (Jee, 2019).

In conclusion, PBA in speaking and listening skills represents an important change of paradigm in language teaching. This assessment approach, supported by specific tasks, offers a comprehensive framework that includes collaboration and the use of technology. To develop assessment literacy to carry out this process effectively, teachers must develop assessment literacy (Firoozi et al., 2019; Prastikawati et al., 2022).

1.1.3. Integrated Skills Testing

Integrated skills assessment refers to methods that evaluate multiple language skills simultaneously. Such an approach, which is especially important in language teaching, requires students to use various skills—reading, writing, listening, and speaking—in an integrated manner.

It is based on the understanding that language skills are interrelated rather than independent, and therefore should be assessed together in authentic, real-life contexts (Brown, 2004; Fulcher, 2010).

There are various methods developed based on this understanding. One of these is Task-Based Language Assessment (TBLA). According to Ellis (2003), this approach involves students completing meaningful tasks that require them to use multiple skills simultaneously. Tasks such as role-plays, group work, or project-based activities allow students to observe their skills and improvements more comprehensively. Portfolios, which contain students' work over a period of time, offer the opportunity to assess not only the final products but also the process and students' reflections (Barrett, 1994). Similarly, PBA methods require students to apply their skills in real or simulated environments (Bachman and Palmer, 1996). Integrated assessment of multiple skills provides several educational advantages. First, since individuals in real life need to use more than one skill simultaneously, such assessments reflect the reality they experience and increase students' motivation (Council of Europe, 2001). Second, teachers can gain a more complete view of students' strengths and areas for improvement (Weir, 2005), allowing the teaching process to be customized for each student individually. Third, this type of assessment encourages students to analyze and apply knowledge across different tasks and contexts, thereby contributing to learning persistence and depth of understanding (Biggs and Tang, 2011).

However, despite all these advantages, integrated skills assessments come with some challenges. Assessing more than one skill simultaneously can cause various difficulties in terms of reliability and validity. In order to ensure fair and balanced assessment of each targeted skill, tasks need to be carefully designed, and teachers need to receive adequate training (Alderson and Wall, 1993). Additionally, such assessment practices can be time-consuming and resource-intensive. Especially in crowded classrooms, the individual feedback process poses a significant challenge (Hughes, 2003). Furthermore, integrated tasks need to be designed in a culturally neutral way; otherwise, injustice may arise for some student groups (Norris, 2002).

In recent years, technological developments have introduced new possibilities for integrated skills assessments. Digital portfolios and electronic assessment tools provide students with more flexible and accessible assessment environments. Platforms such as Google Classroom and various Learning Management Systems (LMS) create interactive assessment environments where multiple skills can be measured simultaneously (Baker, 2020). On the other hand, artificial intelligence-supported systems are developing assessment tools that can be tailored to student performance and provide personalized feedback (Shute, 2015). As a result, integrated skills assessments stand out as one of the emerging contemporary assessment approaches that reflect student competencies more realistically.

1.2. Assessment

The term "test" emerged in the early twentieth century, when exams were employed to choose candidates for university admission. These assessments were initially oral, then written, and targeted certain circumstances and goals. The term "assessment" has since come to refer to all the activities that teachers do to help students learn and track their progress. In education, assessment is now regarded as a learner-centered and learner-oriented goal (Tsagari and Banerjee, 2016). In the past, "language testing" was used instead of "assessment," and assessment has been used more recently.

Black and William (1998, p. 143) define assessment as "all activities teachers use to help students learn and gauge student progress". Assessments serve a different purpose than tests, focusing on the learner. However, its meaning has gradually evolved to something more akin to a kinder, softer type of assessment, and it can now be applied not only to tests, but also to any other activity that allows one to learn about someone's capabilities. Two things need to be distinguished in the purpose of assessment: achievement and proficiency. Language achievement exams assess specific course topics, whereas proficiency tests assess general language skills regardless of the teaching sequence content.

1.3. Types of Assessment

1.3.1. Summative and Formative Assessment

Scriven (1967) introduced the concept of formative assessments in education, which emphasizes the use of ongoing assessment techniques to shape learning outcomes. In this position, it might promote learning rather than just determining the final grades (as in the summative evaluation). The fundamental goal was to promote student responsibility through motivation, teacher feedback, and self- and peer assessment.

The concepts of "formative" and "summative" assessment do not appear to be difficult to understand; however, their definitions have often been misinterpreted in recent years, particularly in the case of formative assessment. Both summative and formative evaluations are used to obtain information within a balanced assessment system. If there is heavy reliance on one method, the reality of student achievement in the classroom becomes unclear.

Summative assessments are given regularly to determine what students know and do not know at any one time. Summative assessments are employed in district and school programs, not just standardized tests like state assessments. Summative evaluation at the district/classroom level is a general accountability measure (Garrison and Ehringhaus, 2011).

In formative assessment, researchers emphasize that focus is an integral part of the teaching process. When integrated into classroom practice, it gives the knowledge required to alter teaching and learning as they occur. In this manner, formative assessment informs both teachers and students of student knowledge, allowing for timely improvements. These changes help ensure that students meet specific, standards-based learning objectives within a specified time limit. Although formative assessment procedures can take many different forms, there are several key differences between them and summative examinations. One distinction is to view formative assessment as "practice." Teachers do not hold students responsible in "grade book fashion" for abilities and concepts that have been recently introduced or are currently being acquired. Therefore, one must allow for practice. Formative assessment helps teachers determine the following stages in the learning process, as it informs teaching approaches and complements summative assessment of student learning.

Another key feature of formative assessment is the active participation of students. Formative assessment cannot be effectively implemented unless students are engaged in the assessment process. Providing students with detailed feedback during learning is an essential component of promoting self-assessment. Research indicates that descriptive feedback is one of the most effective instructional strategies for supporting student progress (Garrison and Ehringhaus, 2011). Therefore, teachers are encouraged to design an integrated approach that combines both formative and summative assessments to gain a comprehensive understanding of students' learning and enhance their learning outcomes.

1.3.2. Informal and Formative Assessment

1.3.3. Classroom Assessment

In the 1990s, textbooks for pre-service and in-service teachers started to emerge that viewed classroom-based assessment (CBA) as a distinct paradigm (Genesee and Upshur, 1996). Over time, research in these areas began to appear in the language testing literature. Well-known online language testing resources have started to include references to CBA (Fulcher, 2010). The assessment of teaching and learning has gained popularity among researchers, and the teacher's role has been emphasized.

Colby-Kelly and Turner (2007) collaborated to define the characteristics of CBA. It involves teachers developing strategies for planning and carrying out the collection of various forms of information about student language use, analyzing and interpreting it, providing feedback, and using this information to assist in making decisions to improve teaching and learning. Various methods (e.g., observation, portfolios, conferences, journals, surveys, interviews, projects, task sheets, and quizzes/tests) are used to collect visible evidence of learning (or lack thereof), which is

typically integrated into regular instructional activities. In other words, CBA is a set of methods and reflection procedures that teachers and students use to continuously assess student learning. This approach adapts instruction to the needs of students (Fulcher and Davidson, 2013).

According to Fulcher and Davidson (2013), various terms are used to describe classroom assessment practices. These include alternative (Fox, 2008), authentic (O'Malley and Pierce, 1996), dynamic (Lantolf and Poehner, 2008), diagnostic (Alderson and Huhta, 2005), performance-based (McNamara, 1997), classroom-based (Genesee and Upshur, 1996), teacher-centered (Davison and Leung, 2009), school-centered (Davison, 2007), and assessment for learning (Black et al., 2003). Today, these terms are often used interchangeably with formative assessment (Brookhart, 2007).

1.4. Assessment Literacy and Assessment Literacy in Türkiye

Assessment literacy refers to a teacher's understanding of how to evaluate what students have learned and how to utilize the data gathered from assessments to enhance students' learning and instructional methods (Pastore, 2023). In addition to theoretical knowledge, a teacher should understand how to guarantee the reliability and validity of assessment tools, as well as the techniques and ideas that impact the assessment process. In practice, teachers work collaboratively with their colleagues to develop evaluation procedures that directly assess what students have learned (Braney, 2011).

Assessment-literate teachers need to understand both theoretical and practical aspects of assessment and evaluation (Yastıbaş and Takkaç, 2018). Furthermore, such literacy indicates how far a teacher may reflect their evaluation expertise in various ways and in diverse contexts. Fulcher (2012) defined assessment literacy based on a study of language teachers' needs in assessment practice. He claimed that assessment literacy has three components. The first is the ability to design, improve, preserve, or assess standardized tests. The second is understanding the assessment method, concepts, and principles that influence practice. The final one is the ability to integrate knowledge, processes, and theories within a social and philosophical framework to comprehend why assessment procedures exist, the role of assessment, and how assessment affects classrooms and individuals. Newfields (2006) takes a different approach to defining evaluation literacy. He classifies assessment literacy from the perspectives of a professional test developer, university student, and high school instructor. Professional test developers perform all of their work with assessment literacy in mind.

It is widely recognized that teachers need to understand and effectively employ all types of assessment purposes (Volante and Fazio, 2007). Assessment literacy is crucial because it influences many aspects of teaching, including decision-making for large-group instruction and the development of instructional programs (Mertler, 2003). Therefore, teachers must possess

competence in assessment literacy to shape instruction and enhance student learning (White, 2009). Stiggins (1991) also emphasizes that teachers should collect sufficient classroom assessment data to make informed judgments about students' learning progress through effective instruction and ultimately to maximize the benefits of their teaching decisions.

Language evaluation literacy has also become increasingly important, given the growing body of literature on the subject (Yastıbaş and Takkaç, 2018). Language assessment literacy is in great demand among instructors, as evidenced by numerous studies (Fulcher, 2012; Ölmezer-Öztürk and Aydin, 2018). Various scholars have conducted studies on different elements of the field worldwide. In the international arena, some research has focused on instructors' assessment knowledge and its impact on assessment procedures. According to Kiomrs et al. (2011), Iranian EFL teachers have limited expertise in language assessment; hence, they rely on standardized examinations as their sole assessment source. Their tests were identical copies of standardized tests with or without change. According to the researchers, participants were unable to resolve the washback effect of standardized tests on students because of a lack of assessment knowledge. In a comparable study, Xu and Brown (2017) discovered that language assessment literacy among Chinese university teachers is low; initial and ongoing teacher training is insufficient; thus, they employ standardized exams. Other studies have focused on teacher training. Tsagari and Vogt (2017) investigated teachers' perceived level of language assessment literacy and their language assessment training needs in seven European countries. It was concluded that language teachers did not have enough assessment literacy training. The issue of insufficient training causes teachers to conduct assessment instruments inappropriately in their classes and to gain undesirable experiences.

According to Kahl et al. (2012, p. 8), "assessment literacy includes the ability to pick and develop many purposes, such as accountability, educational program evaluation, student growth monitoring, and promoting and identifying unique student needs". Actually, pre-service programs must be the main subject of assessment literacy analysis in order to determine their suitability for training candidate teachers (Öz and Atay, 2017). According to Malone (2017, p. 225), "...there is no consensus on what is required or even needed for language instructors to reliably and validly develop, select, administer, and interpret tests". Furthermore, Shepard (2000) claims that "typical educational measurement courses taught pre-service instructors domain requirements, item formats, and methodologies for measuring reliability and validity, resulting in inefficiency." Few connections were provided in subject matter methods courses to demonstrate how testing may be used to inform instruction (Shepard, 2000, p. 1). In her paper, Brookhart (2001) asserts that teachers are ill-equipped to assess standardized tests and use appropriate grading strategies. According to Stiggins (2001), a significant number of teachers and administrators are not assessment literate, which leads to students being misassessed and unable to realize their full potential (Mertler and Campbell, 2005, p. 4). Muñoz et al. (2012) investigated whether the

assessment procedures and beliefs of 62 Colombian instructors aligned. They employed interviews, written reports, and surveys as instruments. The survey revealed a discrepancy between the opinions and approaches of teachers, underscoring the need for improvement.

Mertler (2003) also conducted a study to investigate the assessment literacy of pre-service and in-service teachers. The study involved 67 pre-service teacher candidates who had recently completed their classroom assessment course at the university. Half of the participants were in-service teachers, with 197 working in district schools. The researcher employed the Classroom Assessment Literacy Inventory (CALI), which has 35 components. The findings revealed that typical teacher preparation courses are not realistic and do not adequately prepare teachers for real-world classroom practice. The author stresses the importance of pre-service education in teacher assessment, stating that "they often believe they have not received sufficient training in their undergraduate preparation programs to feel comfortable with their skills in making assessment decisions" (p. 22).

Jannati (2015) studied 18 EFL instructors with various degrees of knowledge (low, mid, and high). Semi-structured interviews were conducted in this study. This study aimed to examine how EFL instructors integrate evaluation literacy into their practices and how experience affects both. The study found no relationship between experience and assessment literacy. Participants who were skilled in evaluation struggled to apply their knowledge to their practice.

Limited research has been conducted in the Turkish context regarding classroom assessment (Hatipoğlu, 2017). Nonetheless, the testing system in Türkiye has changed significantly in recent years (6 times since 2003) (Bilen et al., 2022; Han and Elçiçek, 2021; *Milli Eğitim Bakanlığı Ölçme ve Değerlendirme Yönetmeliği*, 2023; Yıldız et al., 2024), and language teachers must help students succeed in the classroom. Hatipoğlu (2015) studied 124 pre-service teachers at the Middle East Technical University (METU) in Türkiye between 2009 and 2012. The study intended to understand their knowledge of language testing and their expectations for their departmental English Language Testing and Evaluation (ELTE) course, including its content and methodology. The survey found that practically all potential ELT teachers expected to analyze, choose, and write exams, as well as prepare students for national and international examinations, using knowledge gained from the ELTE course. In Türkiye, pre-service teachers' opinions about the ELTE department course at universities are influenced by the local assessment context, which focuses on exams (Hatipoğlu, 2015). Interestingly, even after four years of ELT instruction at METU, the students had minimal knowledge of assessment.

Haznedar (2014) also conducted a study on 538 primary school teachers to see what assessment techniques were utilized to evaluate children's achievement. The study found that most participants (98.9%) used outdated written exams and multiple-choice questions (78.2%) to assess the performance of primary school students. According to Köksal (2004), there is little training in

terms of evaluation, particularly in testing. When the word 'test' is mentioned, most Turkish language teachers immediately think of the 'multiple-choice exam.'

Researchers in Türkiye have also investigated the challenges in assessing listening and speaking. Aşılıoğlu (2009) examined listening issues among 195 Turkish teaching students. Long teaching hours, unclear lesson objectives, student indifference, passive roles, unfamiliar concepts, and ineffective teacher body language were among the barriers. Emiroğlu (2013) also surveyed 97 preservice Turkish language teachers about listening barriers. They identified problems originating from the speaker (58%), listener (21%), and other factors (21%). According to Kaldırım and Degeç (2017), some reasons make it difficult to understand what they hear, and these reasons are speech speed, their limited vocabulary knowledge, dialect issues, as they cannot understand dialects of the speakers, and idioms and proverbs.

1.5. Listening

Listening is a complex process, although it appears to be a passive activity. Vandergrift (1999) defined listening as a passive process that involves analyzing sounds, vocabulary, syntax, stress, and intonation in a specific environment. However, there are certain ways to deal with hearing, and phonological knowledge is based on the differences between learners with different competence levels, as well as the tactics used (Elmankush, 2017).

Nihei (2002) suggests seven listening strategies. One of them is listening for the main idea, which ignores general information questions and details. At the same time, another listening strategy involves listening for specific details, focusing on details to enhance comprehension. Listening to predict is another strategy, including predicting the present and future, decoding rhetorical and paralinguistic elements such as facial and body movements, and finally listening to make inferences, which focuses on understanding the true goal and message. Among these tactics, listening for the main idea, listening for precise details, hearing to predict, listening to make inferences, and using nonverbal clues were considered when developing the test items in this study.

Hamouda (2013) emphasized the importance of listening skills in language education, stating that without them, students cannot improve their speaking abilities. According to Mendelsohn (1994), language learning requires active listening (40-50%), followed by speaking (25-30%), reading (11-16%), and writing (9 -10 %). Listening and speaking are important skills for L1 learners, although foreign language learners may struggle with them more than others (J. Rubin, 1990). Ayuanita (2013) notes that formal and informal variants differ in tone of voice, intonation, rhythm, and background noise. The researcher found that when learners listened to a task multiple times, they were better able to focus on the questions once they understood the scene or circumstance. However, they may require more resources.

Listening is a multifaceted skill that involves both linguistic and non-linguistic aspects (Vandergrift and Tafaghodtari, 2010). Zhu (2011) emphasizes that listening, alongside speaking, stands out as the most crucial skill for communication. Winke et al. (2010) explain that linguistic abilities encompass the analysis of speech units and word meanings through long-term memory, in addition to interpreting non-linguistic cues, such as facial expressions and bodily movements, to acquire pragmatic knowledge. In her view, successful comprehension of a spoken text necessitates the integration of both types of skills. Daniel (2011) notes that once a portion of speech is processed and encoded in working memory, it is associated with long-term memory, which draws upon prior experiences and knowledge of the world. This information is then related to the contextual features of the listening text, with facial expressions and body language serving as illustrative examples. Moreover, figurative language may convey intentional meanings that extend beyond the literal interpretation of words, requiring listeners to discern rhetorical devices such as irony and metaphor. These insights collectively suggest that listeners interpret information based on their prior knowledge, indicating that the listener's own cognitive framework ultimately constructs meaning.

Students should practice listening skills in the classroom just as they would at home, since it is the first language skill they acquire (Woottipong, 2014). Children are typically exposed to verbal input before developing other skills such as speaking, writing, and reading. As listening develops, these skills follow in parallel (Underwood et al., 1989). Effective communication depends on strong listening abilities, and engaging in listening practice supports the improvement of other language skills over time (Krashen, 1989; Woottipong, 2014). Improving listening skills enhances both the reception and comprehension of information. Experienced listeners are better at predicting context and assessing content due to their expanded long-term memory, which encompasses world knowledge and contextual cues. This predictive ability contributes significantly to listening comprehension and develops through accumulated experience. Empathy, cultural understanding, and background knowledge are also vital components of effective listening, although they vary from individual to individual, leading to different interpretations of the exact text (Hedge, 2006).

Listening comprehension involves two basic strategies: bottom-up and top-down. The bottom-up approach emphasizes decoding sound segments and understanding linguistic structures. In contrast, the top-down strategy builds contextual meaning by activating the listener's prior knowledge to grasp the overall message (Vandergrift, 2011). The ability to anticipate content is, therefore, a top-down process. Background knowledge plays a key role in active listening, as learners must assess and integrate incoming information with their existing understanding (Long, 1989). Those with stronger prior knowledge can infer meaning more effectively from target-language cues (Tyler, 2012). Vandergrift (2011) identifies three stages of listening comprehension: perception, parsing, and usage. Perception is linked to bottom-up processing, while usage aligns with top-down strategies—especially for advanced English learners (Osada, 2004). The parsing

stage, on the other hand, can involve both approaches. Lower-level learners are encouraged to focus on content words to derive contextual meaning. In contrast, higher-level learners can employ more holistic, top-down strategies to comprehend the message more effectively (Vandergrift and Tafaghodtari, 2010). This distinction highlights the importance of equipping lower-level learners with appropriate skills to facilitate comprehension.

At this point, it is helpful to summarize Buck's (2001) comprehensive insights on listening. He emphasizes that bottom-up processing entails breaking language input into phonemes that convey meaning, followed by decoding and analyzing words for their basic informational content. This process engages multiple cognitive stages and is linear, mainly, or in Buck's words, "a one-way street" (p. 2). Conversely, listeners also rely on top-down strategies, drawing on general knowledge and communicative intent to interpret meaning, choosing specific words to match social situations and expectations. In addition, Buck distinguishes between transactional and interactional language: while the former has a clear communicative goal (e.g., a business meeting), the latter focuses on the manner of speaking, where how something is said may outweigh what is said. When both forms are used simultaneously, one typically dominates depending on context.

Listening comprehension is also shaped by prosodic features such as stress and intonation, particularly because English tends to highlight key information through these cues. Stressed syllables are usually louder and may be preceded by a pause, making them easier to identify and more likely to carry essential meaning. Understanding these patterns is especially important for learners, as stressed words often carry multiple interpretations (Brown, 1990).

Researchers have also identified other difficulties in learning strategies for listening and comprehension. Chen (2005) studied 64 Taiwanese university students to identify difficulties in learning listening strategies. According to researchers, barriers included affective, habitual, information processing, English proficiency, strategic, belief, and material-related issues. Vandergrift (2006) researched 133 native English-speaking students learning French. He hypothesized that listening comprehension difficulties in foreign language learning stem from the challenges of listening to the native language and insufficient knowledge of the target language. Results showed that 25% of foreign language proficiency and 14% of native language listening skills are predicted by foreign language listening comprehension.

Another issue that may prevent foreign language learners is hesitation, which comprises elements such as unfilled pauses, filled pauses, and repetition (Buck, 2001). Voss (1979) conducted an experiment and determined that one-third of the problems found in comprehension were due to hesitation, in which receivers may lack critical information about how to decode the message or miss a portion of it. Buck (2001) described two types of language use: transactional and interactional. The goal of transactional language is to ensure that communication is successful.

Elmankush (2017) believes that proficiency is measured by assessing learners' vocabulary, phonology, and background knowledge. According to Bloomfield et al. (2010), listeners with a vocabulary of over 5000 items are more likely to understand what is said. Vocabulary plays a crucial role in listening comprehension, as certain words may contain valuable information.

Metacognitive strategies can help learners overcome issues with listening, including misunderstandings and unfamiliar languages (Goh and Vandergrift, 2021; Vandergrift and Goh, 2012). Vandergrift and Tafaghodtari (2010) identified five ways in which learners can improve their understanding. These include planning and evaluation, problem-solving, avoiding mental translation, self-awareness, and direct attention.

Actually, "little is known about the variables that contribute to the development of L2 listening ability" (Vandergrift and Baker, 2015, p. 391). Individual differences in L2 listening skills can be attributed to various linguistic and cognitive factors. As the researcher suggests, individual differences and many variables have an impact on the assessment and acquisition of listening and speaking skills. Therefore, despite many years of learning, students are not as fluent as they are in reading or writing. Metacognitive knowledge (MK) and working memory (WM) are key cognitive characteristics that influence L2 listening performance. Bloomfield et al. (2010) found that L2 listening anxiety, a situation-specific anxiety, could affect L2 listeners' performance. These are just some of the challenges in assessing students' abilities, and there are many more things to consider when planning and applying these skills carefully in a classroom atmosphere. Developing listening skills in a second language involves constant practice and extensive exposure (Rost, 2006).

To summarize, developing and assessing listening skills in a second language is a complex procedure influenced by linguistic, cognitive, and contextual factors. Therefore, teachers need to apply effective strategies considering individual differences, and they should create opportunities for students' meaningful practices for successful results in real-life communication.

1.5.1. Listening Assessment

Listening is a critical component for EFL learners and is challenging to acquire and assess. When literature is reviewed, it becomes clear that listening is not a skill that can be acquired and assessed in a short period. It requires an active process of involvement and cannot be evaluated just by tests. Sanavi and Mohammadi (2018) suggested that listening is more than just receiving spoken ideas; it is the active process of creating meaning through aural stimuli, followed by additional processes that eventually lead to comprehension. In doing so, the listener must go through a series of steps, beginning with the processing of raw speech. The next step is to determine the type of speech event the listener is exposed to, as well as the function of the speech act. The listener activates schemas based on the context and topic matter to perceive the message effectively.

According to Sanavi and Mohammadi (2018), another step in the process is to provide literal meaning to aural inputs, which is followed by the intended meaning. This stage is characterized by the listener's semantic interpretations of what they hear. A variety of factors, such as the listener's lack of interest, conceptual/linguistic complexity, or even the rate of delivery, might hinder listeners' comprehension of the message being conveyed, making the utterance useless. Next, the listener decides whether to commit the utterance to short- or long-term memory, after which the form is deleted, while the message's "truth value" remains in memory. Another challenge is, for instance, that the assessment process for EFL learners can be "selected" responses for listeners. However, a listening assessment needs to be well-designed and based on critical thinking about many things. Chen (2004) argues that selected responses require test takers to recognize and choose an appropriate response, whereas constructed responses involve recalling aural information and creating their own responses. According to Krashen (2002), one of the most essential functions of classroom listening activities is to prepare students to deal with natural and communicative circumstances.

Learners are not prepared for spontaneous, day-to-day interactions with native speakers unless they are able to decode new messages by using familiar terminology and structures. To determine whether learners possess sufficient language skills to handle everyday encounters with native speakers, teachers rely on learners' reconstructions of meaning in comprehension assessment tasks to predict how effectively they comprehend auditory information in real-life listening situations. Consequently, the evaluation task needs to be relevant to and consistent with what people typically do in their daily communication.

1.6. Speaking

Speaking is one of the most challenging language skills to measure since it is dynamic and interactive, necessitating live observation or recording for later evaluation (Ginther, 2012). Speaking is also challenging for students; therefore, students' speaking performance could be much more challenging for EFL and ESL teachers to assess. Alfa (2020) suggests that the primary challenge in acquiring and evaluating speaking skills stems from cultural factors, specifically the students' perceptions and comprehension of the language they are learning. In some cultures, students may find it unusual to participate actively. Learners feel that their function in the classroom is to maintain silence and listen to teachers' explanations. They are bound by feelings of reluctance to participate in class. According to the researcher, the second category is linguistic factors. Linguistic factors refer to word choice, spelling, and verb tenses. The third aspect is social and refers to their capacity to communicate with others in English. In other words, some learners have little practice in utilizing English. The challenges faced by EFL students in speaking classes stem from both students and teachers. As a result, teachers and policymakers must be able to identify the root cause of problems in speaking lessons to find solutions.

Rajendran and Yunus (2021) also imply that speaking, which is regarded as the most important skill (Richards, 2008), remains the most challenging skill to learn among other language skills (Bailey and Savage, 1994), and is even more difficult for ESL and EFL learners worldwide (Abugohar et al., 2019; Leong and Ahmadi, 2017; Paneerselvam and Mohamad, 2019). Abugohar et al. (2019) stated that EFL learners often view their inability to speak competently as an actual failure in language learning. Furthermore, speaking abilities are considered crucial for the future job prospects of ESL and EFL learners (Rao, 2019), and being unable to communicate in English is one of the primary causes of unemployment. Thus, acquisition of speaking abilities has become critical for both ESL and EFL learners, and there is a need for a study of appropriate pedagogical strategies, tactics, and approaches that can illuminate teachers on how to make speaking teaching and learning more successful.

Learners in the classroom generally use English for learning goals, and when they have difficulty speaking in English, the entire educational process becomes difficult to carry out and achieve success. Despite English being a required subject in primary education and emphasizing conversational skills, students' inability to communicate in English is problematic (Islam et al., 2022). Another reason for the challenges in speaking and listening may be a lack of exposure to the four skills among students, as they are already proficient in these skills in their mother tongue. According to Akram and Malik (2010), in real life, people communicate in multiple languages simultaneously. The integration of these four skills stresses the importance of realistic language and can assist language teachers in developing communicative competence among English learners. Integration enables teachers to include various activities in their lessons.

Speaking is a crucial component of communicative language, as it facilitates the integration of other language skills. According to Leong and Ahmadi (2017), speaking enhances vocabulary and grammar, which in turn contributes to improved writing skills. Learners who can express their ideas, tell stories, and engage in discussions orally find it easier to articulate their thoughts in writing. Research also indicates that speaking and reading skills are closely interconnected (Richards, 2008). In other words, as learners' speaking skills improve, their reading proficiency tends to develop as well. Krashen (1989) further highlights the strong relationship between speaking and listening skills, arguing that learners demonstrate their language acquisition through successful listening and speaking performance.

Additionally, practical communication skills may help with socialization (Lucanus, 2017). Socialization occurs when an individual constructs a message, communicates it, and interacts with others. According to Paneerselvam and Mohamad (2019), 'social skills' are advantageous as they foster connections between individuals. 21st-century talents include collaboration, problem-solving, critical thinking, and effective communication. These are some of the attributes that employees are

expected to have in today's globalized environment. Rao (2019) supports this by stating that having good communication skills is a passport that improves work opportunities.

Furthermore, speaking involves using language to convey ideas, feelings, and thoughts. Acquiring this language ability is essential for effective communication with others. Furthermore, speaking is an activity that plays a crucial part in classroom language learning. Speaking allows students to express their creativity by expressing their ideas and views. Barnes (2003) supported the notion that talking is vital for sharing ideas and communicating in the classroom. Speaking helps students convince, inform, and raise questions from their own perspectives.

During teaching and learning, teachers must consider several aspects of speech (Duong, 2014). Moreover, Bustari et al. (2017) suggested that grammar, fluency, correctness, vocabulary, and pronunciation are the most significant parts of speaking. The first component is vocabulary. Vocabulary refers to words that serve multiple functions. Words clarify communication and prevent misunderstandings. It helps to define many objects, activities, and thoughts. Vocabulary mastery indicates linguistic comprehension skills. The second component is grammar. This is akin to strict restrictions on both spoken and written forms. Grammar is a set of principles that control the standard order and connection of words in a phrase. The third category is pronunciation. This refers to the correct pronunciation of a word. Proper pronunciation is important for effective communication. Pronunciation is important in speaking, as it conveys the intended meaning of what is being said. The fourth characteristic is fluency. It entails understanding how he or she delivers thoughts without utilizing filler words ("um" and "ah") while conversing with the other person. According to Yang (2014), fluency is the capacity to talk at an effective speed, smoothly, and with proper words without overthinking. In this scenario, excellent communication skills boost both confidence in presenting ideas and communication abilities.

The fifth quality is accuracy. Kusnierek (2015) defined accuracy as the ability to understand words, structures, grammar, and meaning without errors. It also relates to the sequential framework of a speaking event, also known as a genre (Samad and Adnan, 2015). Teachers must examine these five elements when teaching speaking skills. Both the student and the teacher should possess these features. These should be stressed for students to achieve a strong speaking performance. To acquire these characteristics, EFL students should be provided with authentic material. They can learn from content written correctly in English. As a result, these qualities must be taught in EFL classes. Teachers must consider these elements when teaching speaking skills. Teachers should present real materials to aid in the teaching and learning process.

1.6.1. Speaking Assessment

Speaking assessment may require people to be equipped with linguistic, vocabulary, grammar, functional, strategic sociolinguistic, and cultural competence, which depends on the context in which the assessment of the speaker is done. This theory is called the criteria for speaking assessment (Rahmawati and Ertin, 2014). A speaking assessment may also depend on what should be assessed in grammar, pronunciation, task, fluency, vocabulary, and comprehension.

Additionally, some factors need to be considered when assessing speaking. These are practicality, validity, reliability, and authenticity. For instance, if one assesses speaking ability, making an essay is not a valid test, and the test designer needs to know this. In addition, rubrics and scoring criteria are essential, and designers need to prepare a standard system before conducting the speaking test directly. The items need to represent all the things that are intended to be measured, and the test takers must be informed clearly of the weight of the items. These are crucial for reliability, and last but not least, one is authenticity, which focuses on giving importance to students' values as performances.

Another aspect of speaking assessment is that all assessments are conducted for some reasons, such as formal and informal assessments. Sheetz et al. (2018) imply that formal and informal assessments are used for placement in education and workplaces, with the choice of assessment method based on factors such as the speaker's level, purpose, time, and skills. Brown and Abeywickrama (2010) provide a taxonomy for speaking tasks, including imitation, intensive, responsive, interactive, and extensive. Tasks should match the target range and require minimal interaction between the speaker and assessor.

Sheetz et al. (2018) argued that to assess a speaking task, it is necessary to listen to speech production samples to gain a reliable grasp of indications such as presence, correlation, and other speech-specific characteristics. It is also critical to consider environmental factors, such as previous performance, other speakers' performances, and the role of the examiner. Speaking is an interactive communication strategy that emphasizes conversation and discussion. The frequency of evaluations is critical, as speaking tasks have limitations when compared to writing, listening, and reading exams. Single assessments can consume instructional time and raise issues regarding what to do with disengaged students. Planning the quantity and type of assessments within the assessment cycle is critical, as is creating a calendar with assessment dates and times.

A successful communication class integrates four language skills and requires the instructor to create a good environment, provide suitable activities, motivate learners, and deal with concerns carefully (Davies and Pearse, 2000). For these reasons, students need to be exposed to four skills, and teachers need to be educated to actively participate in students' communicative skills

throughout an extended period; if not, assessment cannot be valid, and reflection of real speaking and listening outcomes of the students.

A review of studies on integrated test tasks revealed that the majority have focused on assessing writing skills, with significantly less emphasis on speaking skills. Research on integrated speaking assessment has focused on factors other than topical knowledge, such as test-taking processes and performance (Barkaoui et al., 2013; Crossley et al., 2014; Lee, 2006; Swain et al., 2009). Others, although emphasizing the importance of content, did not investigate whether topical knowledge influences how test-takers obtain content correctness.

Brown et al. (2005) found that raters emphasized content when judging performance on the alone and integrated speaking test activities for the new TOEFL project. (Frost et al., 2011) analyzed the integrated listening-speaking task of an Oxford English language test and concluded that it is feasible to include a content dimension in the speaking assessment and that accuracy of content should be considered in the construct of speaking ability (p. 366). Jamieson and Poonpon (2013) created three analytic rating guidelines for the integrated speaking tasks on the TOEFL IBT, providing teachers and test-takers with more accessible rubrics.

Taken together, these studies highlight the critical role of content accuracy in integrated speaking exam performance; however, it is unclear whether and how test-takers' topical knowledge influences content accuracy. To summarize, practical speaking assessment requires careful consideration of linguistic and contextual criteria, as well as practicality, validity, reliability, and authenticity. Additionally, it involves integrating speaking tasks with content accuracy, topical knowledge, and communicative competence to ensure a comprehensive evaluation of learners' abilities. The purpose of communicative language instruction is to help students develop communicative competence, which encompasses both language knowledge and the ability to use language effectively in communicative contexts (Xu, 2018). A language teacher's goal is to educate students on how to use a foreign language in real-world contexts. However, the present ELT techniques are not in line with this ideal. Therefore, integrated language training and assessment may help bridge the gap between isolated language skills and real-life use.

The literature has shown a clear shift from discrete-point tests to communicative, performance-based, and integrated skill assessments since the 1970s, parallel to the student-centered transformation in language teaching. Instead of exemplifying grammar and vocabulary knowledge in isolation, contemporary approaches seek evidence of how learners simultaneously mobilize multiple skills to achieve real-life goals; task-based practices (role-plays, presentations, projects) and classroom-based assessment routines become widespread. In this approach, assessment is used as a formative tool. Feedback, student participation, and repeated task cycles increase student motivation and support learning. Technology, especially artificial intelligence applications, offers new opportunities to design original tasks and create digital portfolios with

personalized feedback. However, in such assessments, attention must be paid to design criteria such as validity, reliability, feasibility, fairness, and cultural neutrality. Factors such as interaction, simultaneity, and context sensitivity, especially in speaking and listening skills, make the measurement process more complex.

The second main axis is assessment literacy. Effective task design, rubric development, assessor calibration, and evidence-based decisions rely on teachers' theoretical and practical competencies. Studies conducted in different contexts, including Turkey, show that gaps in pre-service and in-service preparation lead to over-reliance on standardized tests/open-ended product-focused assessments, weak alignment between learning objectives and criteria, and limited use of formative evidence. In listening and speaking, teachers must deal with cognitive and affective factors (working memory, anxiety, hesitation phenomena), multimodal cues (prosody, nonverbal indicators), and processing pathways (bottom-up/top-down); which directly affect task difficulty and scoring. These complexities increase the need for robust rubrics, evaluator training, and task banks that enhance consistency while preserving authenticity.

Overall, the literature proposes an action-oriented agenda: transitioning from discrete skills to integrated and communicative tasks; embedding assessment formatively within instruction; and strengthening the capacity to design, implement, and interpret high-quality assessment. The most urgent priorities for the Turkish middle school context, where speaking and listening assessment has recently become mandatory, are targeted professional development (assessment literacy, assessor calibration, task design), structural supports (sample tasks, analytical rubrics, moderation processes), and measured use of technology that increases originality without sacrificing reliability. Meeting these needs will bridge the gap between research, policy, and classroom practice, enabling assessments that both reflect and advance learners' communicative competence.

CHAPTER TWO

2. METHODOLOGY

In this chapter, the research design, data collection tools, and analysis procedures are explained. The study adopts a mixed-method case study approach to explore secondary school English teachers' perceptions of the integration of listening and speaking assessment into the 2023–2024 national curriculum in Türkiye. Both quantitative and qualitative data were collected through questionnaires and semi-structured interviews to gain a comprehensive understanding of teachers' experiences, practices, and challenges. The methodological framework also includes procedures to ensure validity, reliability, and ethical standards throughout the research process.

2.1. Research Design (Case study/mixed method)

This study is a case study that explores the perspectives of a specific group: secondary school English teachers in Ortahisar, Trabzon. It focuses on a current issue—the changes made to Türkiye's national education program during the 2023–2024 academic year. In particular, the research examines how the integration of “speaking and listening assessment” into the education system is perceived by these teachers. According to Stake (1995), a case study investigates the uniqueness and complexity of a single case. In this context, the case is defined by the combination of a particular group of teachers, a specific educational reform, and a defined time frame.

Case study researchers often employ a range of data collection methods, including interviews, observations, and document archives. Although case studies are commonly referred to as qualitative research, they often incorporate quantitative data-gathering methods, such as questionnaires (Verschuren, 2003). Guetterman and Fetters (2018) also note that case studies are increasingly being employed in mixed-method research, such as program assessments.

In this study, a descriptive case study method was adopted within a mixed-methods research design. Both qualitative and quantitative approaches were employed to examine and describe how teachers perceive and experience the recently implemented educational programs aimed at assessing the listening and speaking skills of secondary school students. According to Bailey (2008), experienced researchers first determine the population before selecting the study group, whereas less experienced researchers begin by identifying the smallest feasible sample size before defining the population (Cohen et al., 2002). Following the latter approach, the study first identified the entire population of secondary school English teachers in the first service area in Trabzon.

According to the data obtained from the Directorate of National Education, there are a total of 352 English teachers in the region, 158 of whom teach at the secondary school level. Among these, 80 teachers (50.6%) participated in the questionnaire phase of the study.

Although it may be manageable in size, since this is a case study, it was considered more feasible to use semi-structured interviews as a qualitative component of the study to improve the validity and credibility of the research through triangulation. Additionally, as Dörnyei (2007) suggests, the case study is not a strategy but rather a means of collecting and organizing facts to deepen our understanding of the social being or object under investigation. Therefore, this study aimed to gain a deeper understanding of the incorporation of a new education system through a mixed-methods approach and a case study research design.

Sandelowski (2003) states that there are two main goals for combining methods: to better understand a particular phenomenon as a whole and to compare one set of results with another. This is a mixed-methods case study because it best meets the research objective, which is to gain a comprehensive understanding of a complex subject by examining it from multiple perspectives. This study is also an exploratory case study because it focuses on a new issue, as speaking and listening assessments have recently been mandated in Türkiye's secondary schools, and there is a gap in research on this topic. The study aims to identify thematic patterns related to this issue through thematic coding as part of the qualitative analysis, identifying and analyzing codes and themes regarding teachers' perspectives, experiences, and challenges faced during the new implementation process. It seeks to understand better how teachers apply assessments in real classroom settings. Additionally, the survey component focuses on the quantitative aspect, statistically analyzing numerical data on teachers' assessment practices, their perspectives on the process, and the challenges they encounter in evaluating the speaking and listening skills of secondary school students.

Creswell (2009, p. 14) identified three main mixed-methods strategies: transformational, sequential, and concurrent. By collecting both forms of data simultaneously and integrating them into the interpretation of the overall findings, this study employed a concurrent strategy to provide a comprehensive analysis of the research problem. To achieve the study's aim, a mixed-methods case study approach was adopted. Data were collected concurrently from secondary school English teachers in the First Service Area in Trabzon regarding their perspectives on integrating listening and speaking assessment into the curriculum. Both questionnaires and semi-structured interviews were used to ensure a thorough and multidimensional analysis of the data.

2.2. Quantitative Research

2.2.1. Participants and Setting

Quantitative data collection was conducted using an online questionnaire with 80 Secondary School Teachers in Ortahisar, Trabzon. The language of the questionnaire was Turkish to minimize data loss. For that reason, I administered the questionnaire in Turkish, which is the native language of the participants. Participants were chosen by snowball sampling. The reason of choosing snowball sampling was to reach experienced Secondary School Teachers which is quite impossible with formal institutions and target group was not a random one but a specific professional group. We had the aim of gaining perceptions of the professionals not the aim of statistical representativeness, therefore, snowball sampling was suitable for this research. The questionnaire was administered to the participants via emails, WhatsApp groups and participants shared the questionnaire with their other colleagues who met the criteria. Snowball sampling provided the opportunity to reach a wide number of people from Secondary Schools education setting. Two secondary school teachers suitable for the criteria of being secondary school English teachers dealing with the newly mandated issue of assessing listening and speaking processes were piloted through e-mail and provided feedback to me on items that were inconvenient or needed to be changed and revised. After their feedback, the questionnaire was reorganized after the last overview of the supervisor, and they were asked to identify further members for “chain reaction”. As Dörnyei (2007) suggests, “chain reaction” is required and useful when groups of people are not easy to identify or access to them is difficult. For this purpose, the interview link was transmitted through a Google Form link via a chain reaction by the participants who identified further members in Secondary Schools and were willing to participate. In addition, 34 listening and 34 speaking questions were asked of the interviewees. However, 34 questions were adapted to each ability separately.

In this study, it was difficult to reach all the possible, convenient, and willing Secondary school English teachers in the First Service Area. For that reason, participants were asked to refer to their other colleagues who were involved in assessing the listening and speaking performances of secondary school students in their schools. This approach was designed to engage individuals with direct experience in research study. The sample size was decided initially by obtaining information from the Ortahisar Directorate of National Education. We gathered questionnaires from 80 Secondary School English teachers, which may be a reasonably sample size, as this number accounts for 50.1% of the entire population in First Service Area. Actually, representativeness of sample size was not the focus of this study. Our focus was to achieve the qualitative and quantitative results of teachers’ perceptions on the integration of assessment of listening and speaking complement to each other. Our aim was not to make a generalization but have a deeper insight on teachers’s perceptions in this newly implementation process. These teachers are all from

different schools of the First Service Area, besides having a variety of teaching experiences, features such as age and gender, so that the generalizability of the research was achieved.

Table 1. Demographic Information about Participants

		Frequency	%
Gender	Male	16	20
	Female	64	80
Age (Average age= 38.2, ranging; from 24 to 60)	24-30	10	12,5
	31-35	20	25
	36-40	26	32,5
	41-45	17	21,25
	46-50	6	7,5
	51-60	1	1,25
Education	Bachelor's degree	71	88,75
	Postgraduate	9	11,25
	Doctor's degree	0	0
The BA Programme Graduated	English Language Teaching (ELT)	52	65
	English Language and Literature	24	30
	American Language and Literature	1	1,25
	Translation and Interpreting	1	1,25
	Other	2	2,5
Year of Teaching Experience	less than 1 year	1	1,25
	1-3 years	6	7,5
	4-7 years	8	10
	8-11 years	14	17,5
	12-15 years	22	27,5
	16-19 years	18	22,5
	more than 20 years	11	13,75

The research study was conducted in response to the timely issue of integrating speaking and listening skills into the Secondary School curriculum for the 2023-2024 academic year. The study specifically aimed to address this pertinent issue within a defined population. As the demographic information about the participants suggests, the age range (24–68) and the average age of 38.2 indicate that the study includes teachers with varying levels of experience. The distribution is

reasonably balanced, with most instructors (78.75%) aged 31-45 years old, indicating a reasonable mix of early-career, mid-career, and experienced teachers.

The setting of the study was Ortahisar/Trabzon (accessible) because it had some advantages, such as accessibility, feasibility, logistical convenience, and real-world applicability. The target population was available, feasible, and easily accessible for the research focus. Additionally, follow-ups were required during the interview process; therefore, the setting was convenient for face-to-face data collection. In addition, it was a case study focusing on a relevant and timely issue, so Ortahisar/Trabzon could provide a local but meaningful study for the research. In addition, the meanings deduced from this district can be applied to other districts.

2.3. Qualitative Research

2.3.1. Participants and Setting

Interviewees were selected with purposive sampling; participants were secondary school English teachers who had firsthand experience in a recently mandated event of integration of listening and speaking assessment into the education curriculum in the 2023-2024 period. Therefore, these secondary school teachers were intentionally chosen to collect relevant data in line with the study's objectives. As (Dörnyei, 2007) implies, despite quantitative sampling that necessitates a sizeable sample, qualitative sampling requires a researcher to choose participants that could clarify and describe human experience; therefore, the researcher did not try to find how experience is distributed in the population but tried to find participants that can provide deep insight into the phenomenon of timely issue "integration of speaking and listening abilities to the curriculum of secondary schools" and this goal could be best achieved with purposive sampling that aims this. Six secondary school English teachers from different schools were interviewed in their schools. The setting was secondary schools in Ortahisar, where these teachers work. Semi-structured interviews were conducted because the researcher and the supervisor had already guided the questions. The details of the semi-structured interview are explained below in Section 3.4. Data Collection Instruments section.

There were eleven open-ended interview questions. First, the purpose of the study was explained to them, and then they were asked to sign the informed consent form to see if they were willing to do it. The teachers signed the consent form allowing their voices to be recorded for transcription purposes. It was necessary because it was crucial to inform the respondents that they would be interviewed as an expert or representative of a group or organization. They were told that their statements would always be kept strictly confidential.

At the beginning of the interview, participants were thoroughly informed about the purpose of the study as well as the procedures for data storage and usage. In a detailed way at the initial part of the interview about the purpose of the study and how the data would be kept and used. They were informed that all the data and recordings would be kept secret and destroyed after the research period was completed. They were assured not to share data with other people. I also explained in detail that English teachers and educational institutions may benefit from this research study, so that the participants might be willing to speak. The audio recordings of the interviewees were processed using the “Voiser Dikte” program within the Microsoft Word Program. The rationale for utilizing this tool was to ensure the retention of all critical information, as it offers a comprehensive transcription from a voice-to-text program.

2.4. Data Collection Instruments

To provide a comprehensive understanding of secondary school teachers’ perceptions regarding the integration of speaking and listening assessment into the curriculum, both quantitative and qualitative data collection tools were employed. A questionnaire and semi-structured interviews were used concurrently to address the research question holistically.

2.4.1. Questionnaire

As a researcher, I aimed to develop questionnaire questions that are relevant to the research questions and objectives. Cohen et al. (2007, p.118) suggest that questionnaires should be transparent about their goals, include questions aligned with research objectives, and incorporate any sub-questions that require empirical data. Based on this principle, two main research questions were formulated, focusing on both listening and speaking abilities. The questionnaire included 34 listening and 34 speaking questions, arranged in parallel. It consisted of multiple-choice questions to facilitate the collection of comparable responses. To achieve this, multiple-choice questions were used to assess teachers' levels of agreement, perceived difficulty, or satisfaction regarding each issue. Additionally, I included open-ended questions to gather qualitative responses that reflect teachers’ detailed opinions. All these question types were designed to be used together in both the listening and speaking sections of the questionnaire, aiming to create a well-structured instrument that gathers information for both the qualitative and quantitative aspects of the study.

Although the thesis is written in English and situated within the field of applied linguistics, the questionnaire was administered in Turkish. This decision was made to ensure clarity, accuracy, and validity in participants’ responses. Although the participants were English teachers, conducting the questionnaire in their native language minimized the risk of misinterpretation, enhanced comprehension, and allowed them to express their ideas more precisely. The purpose of the study was not to measure participants’ English proficiency but to explore their perceptions and

experiences related to assessment practices. Using Turkish, therefore, helped eliminate potential linguistic barriers that might otherwise influence the reliability of the data. This approach aligns with the principle of collecting data in the participants' most comfortable and familiar language to ensure the authenticity and trustworthiness of their responses (Dörnyei and Taguchi, 2010).

The questionnaire was planned to provide answers to the research questions; for instance, some questions were intended to identify assessment practices of teachers (RQ1- RQ2), and there are questions for assessing strategies in the questionnaire as well (S5, S6, etc.). These were explained in detail in the data collection procedure part. The research questions effectively align with the questionnaire and research objectives. The questionnaire had five main parts; the first part was demographic information about participants. The second part focused on teachers' assessment practices, encompassing both abilities. In contrast, the third part focused on their perspectives on the new policy and their ideas for integrating these abilities into the secondary school curriculum. The fourth part deals with challenges and difficulties in implementing these abilities in the classroom for teachers. The last one was about listening and speaking-based questions, the tasks, grading systems that teachers used or applied for these skills, as well as how fair they think the assessment system is that they use. Within these five issues, they were asked specific questions separately in the speaking and listening parts of the questionnaire.

Each skill required detailed and specific questions; therefore, there were a total of 68 questions. However, it is not much, actually, as it was crucial to take into account the challenges, practices, perspectives, assessment procedures, and experiences of the teachers in both skills. For this reason, these questions ensured a valid and comprehensive investigation of the issue, also lending credibility to the research. After piloting with two English teachers in the field and lastly with the supervisor's comment, necessary revisions were made and applied to other participants via Google Form links.

2.4.2. Semi-Structured Interview

For the qualitative part of the research, semi-structured interviews were used. Although there is a pre-prepared set of leading questions, the approach is open-ended and allows the participant to comment freely on the topic. This makes the method ideal for exploring individual situations in depth. The interviewer has a fair grasp of the phenomena and may broaden the questions ahead of time, but he does not want to employ ready-made responses, which may limit the depth of the respondents' stories. As a result, the "interview guide" requires piloting (Dörnyei, 2007). Therefore, I, as a researcher and as a teacher, had already grasped the phenomenon and planned the questions with the aid of the supervisor. Then the interview was piloted with two Secondary School English teachers in a secondary school with face-to-face interviews. After receiving feedback on the questions and the supervisor's support, which drew on my own experience as a teacher in

secondary school, the final version of the interview questions was created. There were 11 interview questions. Even if the sample consists of one example, I must have decided on the features of that case. As a result, a qualitative study requires a sample plan that describes the sampling parameters and is consistent with the study's goal (Dörnyei, 2007). For this purpose, I used purposive sampling to select participants from Secondary School English teachers to hear their voices for achieving the goal of the study related to them. Teachers from different schools were interviewed, and each was informed about the purpose and the importance of the study. They were also informed about why their voice would be recorded and that all the recordings would be kept confidential and then destroyed after the process ended. They also read and signed a consent form approving their participation in the research.

Question 1 and Question 2 were about major research questions asking teachers' general perspective on integrating speaking and listening skills. Questions 3 and 4 were aimed at seeing their idea about possible advantages of this integration and potential impact on student success; their ideas on these issues were aimed at assessing. Questions 5 and 6 in the interview focused on the challenges faced by students and teachers after the integration of speaking and listening skills into the curriculum. Whether teachers have an idea for necessary improvement for these drawbacks, Questions 8, 9, and 10 were about the pedagogical effect of these skills in terms of whether they have motivational, educational outcomes to encourage students, bringing out confidence within students, or not. Finally, the last question was intended to ask teachers' advice for curriculum designers, their feedback on this timely arrangement, and their suggestions for improvement.

The “Voiser Dikte” voice-to-text program was utilized in the Microsoft Office Word program to transcribe the participants' voices into text. In doing so, I aimed to record all information and prevent data loss, which could be possible while taking notes during the interview. After all the details were explained to the participants, their voices were recorded in a silent place.

Thematic coding was applied in this process. After the transcription procedure, I carefully read transcripts and found key words, ideas, and phrases the participants used, and then grouped them into topics that teachers mainly focused on. After that, each theme was defined and analyzed by deducing meaning from the transcripts of the participants. The themes were also thoroughly analyzed in relation to the research questions.

2.5. Data Analysis

A detailed data analysis was conducted to present the teachers' perspectives and experiences in a relevant and accurate manner—the researcher, using mixed methods separately, aimed to achieve a comprehensive data analysis.

2.5.1. Quantitative Data Analysis

Data in the questionnaire were analyzed using SPSS (Statistical Package for the Social Sciences). For this purpose, descriptive statistics were used to analyze demographic information about the participants, assessment practices, and perceptions of secondary school English Teachers in Ortahisar, Trabzon. Descriptive analysis of the 80 participants focused on frequency distributions and percentages to describe variables such as gender, age, teaching experience, and educational background of the teachers. For each question, statistics such as mean, mode, and median values were calculated. These data provided various attitudes of the teachers. Additionally, a comparative analysis was conducted to compare parallel questions related to listening and speaking. The researcher analyzed these parallel questions to see differences and similarities in teachers' perceptions and methods.

I aimed to focus on the challenges teachers face when assessing and other key issues that affect the assessment of listening and speaking, as well as what kind of assessment they prefer or what they suggest doing to improve the assessment of these skills. The results were presented in a detailed manner, accompanied by tables and graphics.

2.5.2. Qualitative Data Analysis

The qualitative part of the analysis employed thematic analysis, which enabled the analysis of data from semi-structured interviews. The interviews were conducted with five English teachers from different secondary schools. The interview was transcribed from voice to text using the Voicer program within the Microsoft Word Program. The researcher then began to read the texts repeatedly to familiarize themselves with significant patterns. After that, phrases, keywords, and issues were carefully coded manually by the researcher. After this process, the key issues, codes, and patterns that teachers repeated were grouped into categories. The identified themes were then analyzed and compared with the research questions and objectives. The researcher interpreted the results, supporting them with statements taken from the participants.

2.6. Triangulation

Data triangulation was a crucial part of the research, serving as both a qualitative component and a complementary part to the interpretation of interview results and questionnaire results. The use of both methods provided an opportunity to ensure the validity of the results because the opinions expressed by the teachers in the questionnaires were consistent with the interview results. Integrating different perspectives from various data collection tools highlighted the drawbacks and strengths of the current assessment procedure, as well as the integration of listening and speaking into the curriculum. Statistical results and content analysis, conducted through thematic coding,

enhanced the reliability and validity of the study. The integration of these skills into Türkiye's secondary school curriculum was evaluated holistically to ensure applicability to future educational plans and curricula.

2.7. Ethical Considerations

This research was conducted in accordance with the principles of scientific ethics and relevant legal regulations. Prior to data collection, ethical approval was obtained from the Ethics Committee of Karadeniz Technical University (Appendix 3: Research Ethics Committee Approval). Additionally, official permission was granted by the Ortahisar Directorate of National Education to conduct the field study (Appendix 4: Research Permission License from the Ortahisar Directorate of National Education).

The confidentiality and anonymity of the participants were strictly maintained, and all data collected were used solely for scientific purposes. No personal information was disclosed to third parties, and privacy was ensured throughout the study. Each participant who contributed to the qualitative part of the research through interviews was assigned an alias (A, B, C, D, E, or F) to protect their identity. Similarly, the expert who evaluated the questionnaire was also identified by an alias to ensure confidentiality. All participants were informed about the purpose, scope, and procedures of the study, and it was clearly stated that participation was entirely voluntary. This information was provided separately for both the quantitative and qualitative phases of the research. In the quantitative phase, participants were informed that submitting the questionnaire (via Google Forms) indicated their consent to participate in the study. In the qualitative phase, participants were asked to sign a written consent form after being fully informed about the study.

Participants were thoroughly informed at the beginning of the questionnaire about the purpose of the study, their right to withdraw at any time, and the intended use of the data collected. It was also clearly explained that the findings of this research could benefit English teachers and educational institutions, encouraging participants to respond conscientiously in order to obtain relevant, valid, and reliable results. Participants were informed that they could contact the researcher via e-mail for any questions or clarifications. Furthermore, the confidentiality and anonymity of the research were explicitly explained. Participants were assured that the findings would not be shared with anyone outside the research process. They were also informed that the results of the study would be presented and discussed at conferences and academic platforms to amplify the voices of secondary school English

In accordance with academic integrity and ethical research standards, artificial intelligence tools such as ChatGPT (OpenAI), Grammarly, and Paperpal were used solely for language editing, grammar correction, and improving the clarity of expression in the thesis writing process.

2.8. Validity and Reliability

Validity and reliability of data collection tools are crucial to ensure the scientific value of the research findings. Validity refers to the extent to which an instrument measures the intended concept (Heale and Twycross, 2015). Within the scope, dimensions such as construct, content, and face validity are considered. Content validity refers to how well the questionnaire items represent all aspects of the construct being measured, in this case, the assessment of listening and speaking skills. Construct validity, on the other hand, ensures that the items accurately reflect the underlying theoretical framework, namely teachers' perceptions, experiences, and practices related to assessment literacy and the integration of listening and speaking assessment into the curriculum. To ensure the validity of the study, expert opinions should be taken, the data collection procedure should be reported in detail, and multiple data sources should be used (Cohen et al., 2007; Yin, 2018; Yıldırım and Şimşek, 2011). Reliability refers to the consistency and replicability of the results produced by an instrument. In quantitative research, reliability is often assessed through Cronbach's Alpha coefficient, while in qualitative studies, reliability is strengthened by expert consultation, supervisor support, and transparent reporting of the data analysis process (Cohen et al., 2007; Heale and Twycross, 2015). In this study, various methods were followed to ensure validity and reliability.

Quantitative data for the study were collected through a questionnaire. To enhance the validity of the questionnaire, the opinions of two different experts were sought. Based on their feedback, six questions that were repetitive or irrelevant were removed. This expert evaluation process directly contributed to the instrument's content validity, confirming that the items comprehensively covered the targeted aspects of listening and speaking assessment. This feedback also improved the questionnaire's clarity, and spelling errors were corrected. Subsequently, a pilot study was conducted with the assistance of two secondary school teachers. Based on the teachers' feedback, some revisions were made. These revisions further supported the construct validity of the questionnaire by aligning the items more closely with the intended theoretical and conceptual dimensions of teachers' assessment practices and perceptions. Through this process of feedback and revision, the questionnaire's validity increased. Cronbach's alpha value for the overall questionnaire was 0.78, while the listening and speaking sub-dimensions were 0.73 and 0.76, respectively. According to Heale and Twycross (2015, p. 77), Cronbach's alpha is expected to be higher than 0.70. Thus, the obtained reliability scores indicate that the questionnaire is sufficiently reliable and consistent to reflect the authentic practices, considerations, and applications of the participant EFL teachers.

Table 2. The Cronbach's Alpha of the Questionnaire, Listening and Speaking Assessment

	Cronbach's Alpha
Overall Questionnaire	0.78
The Listening Sub-dimension	0.73
The Speaking Sub-dimension	0.76

Regarding the qualitative aspect, to improve validity, semi-structured questions are revised based on expert opinions. These opinions help ensure the questions are clear, understandable, suitable for the purpose, and free of unnecessary repetition. Then, pilot interviews are conducted to test the functionality of the semi-structured form. This process also contributes to both content and construct validity in the qualitative phase by confirming that the interview questions adequately represent the intended concepts and remain consistent with the study's theoretical framework. To ensure the reliability of the qualitative data analysis, support was provided by two independent coders who analyzed 20% of the transcripts. Inter-coder agreement was measured using Cohen's Kappa ($\kappa=0.86$). The value indicates strong reliability. To increase data consistency, direct quotes from participants are included, and their opinions are presented as expressed, without alteration. All these steps aim to strengthen the scientific validity of the research and improve the reliability of the findings.

2.9. Positionality in the Study: Being an Insider as a Researcher

In this study, I was both a teacher and researcher investigating EFL teachers' voices about the integration of speaking and listening assessment into national exams in secondary schools in the curriculum. As I have been teaching for many years in secondary schools, my position helped me to interpret the data gathered by my colleagues. It was easier for me to understand the challenges, such as time constraints, workload, and inadequate training problems that teachers face when assessing speaking, especially in formal ways. The insider position of me enabled me to analyze and interpret the data provided by participants, but it also had the disadvantage of the possibility of bias because, as an insider, I may have unconsciously interpreted the data through my own perspective. However, the researcher's positionality or positions are not independent of the research process, nor do they entirely dictate it. This should instead be viewed as a conversation that challenges biases and viewpoints about the social world as well as the researcher. The results and procedure of the research are enhanced by this (Palaganas et al., 2017).

Given the nature of case studies, where the researcher participates in the process and the research context by closely examining the effects of the intervention, positionality emerges as one of the most important factors, particularly in qualitative research, influencing the reliability of

study findings. Since no one can have a completely unbiased position on the topic under investigation and observation, it is nearly impossible to have an absolutely objective point of view, especially in social sciences. Since personal factors such as beliefs, interests, and history reveal our participation with the topic of study, it is vital to recognize their influence on the research context and methodology (Laws et al., 2003).

To lessen the bias in my study, a case study focusing on a specific group, EFL teachers' perception on newly implementation of speaking and listening assessment because of my insider position as a teacher and researcher who was also a part of the target group that investigated, not an outsider that is a stranger to the target group or to the study. To overcome or minimize the insider position, I tried to achieve reflexivity with triangulation of the results, mixed methods, comparing and contrasting the results of qualitative and quantitative data. Besides this, I got help from experts and my supervisor to check especially the themes and codes in the qualitative part of my thesis and my themes were checked by two experts and Cronbach's Alpha was calculated. For the qualitative part, I often questioned my insider position that may affect coding process. The consistency between expert peer review and the codes derived after transcribing the interview data—collected through Microsoft Word Dictate (voice typing)—was confirmed by the calculations of Cohen's Kappa. In fact, for the quantitative part, careful piloting process of the items, expert and peer reviews were also taken to lessen items that reflect my assumptions and increase credibility and neutrality of survey items.

CHAPTER THREE

3. FINDINGS

This study aimed to explore the perspectives of English teachers in the first service area in Trabzon, Türkiye, regarding the inclusion of listening and speaking exams in assessing students' language learning progress. To achieve this aim, both quantitative and qualitative research methods were employed to provide a comprehensive and in-depth analysis. Quantitative data were collected through a questionnaire, while qualitative data were obtained from semi-structured interviews. The quantitative data were analyzed using SPSS software, and the results are presented in tables and figures. Thematic analysis was conducted to interpret the qualitative data.

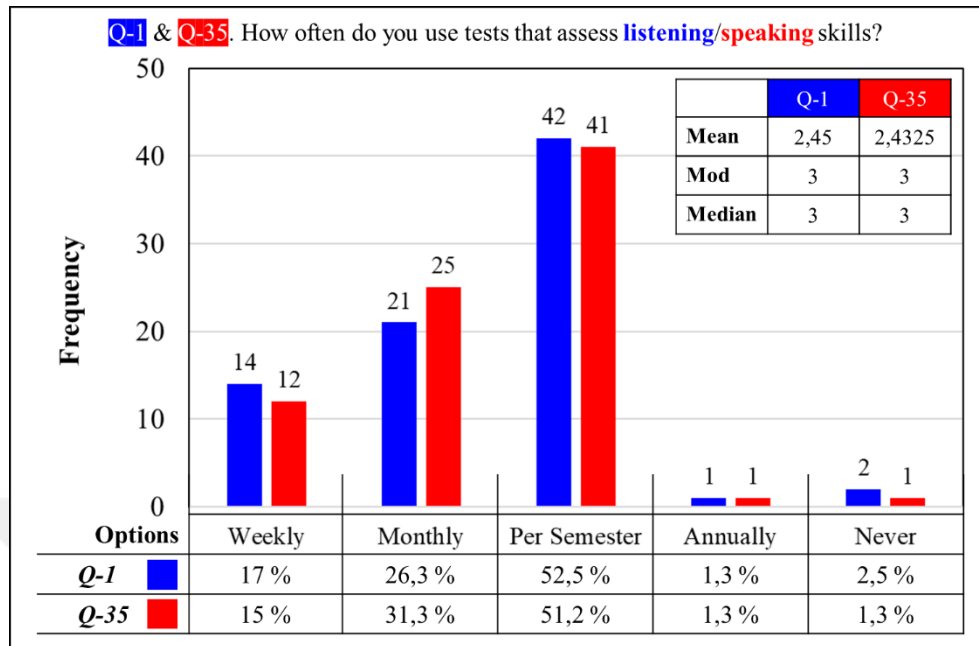
3.1. Quantitative Data

As mentioned in the subsection 2.1 Research Design (Case study/mixed method), this study follows a mixed-method approach; therefore, data were collected via questionnaires and interviews.

The questionnaire began with a section that collected participants' demographic information. The main part of the questionnaire consisted of two sections: *Assessment of Listening Skills* and *Assessment of Speaking Skills*. Each section consisted of 34 questions (Q) that were categorized into five subsections. These sections are organized as follows:

1. Administration and Assessment of Listening/Speaking Exams
2. Reliability and Validity of Listening/Speaking Exams
3. Impact of Listening/Speaking Exams on Teaching and Learning
4. The Use of Feedback by Teachers in Listening/Speaking
5. Teachers' Professional Development and Educational Background on Listening/Speaking

Figure 1: Statistics and Frequency Values for Q-1 and Q-35



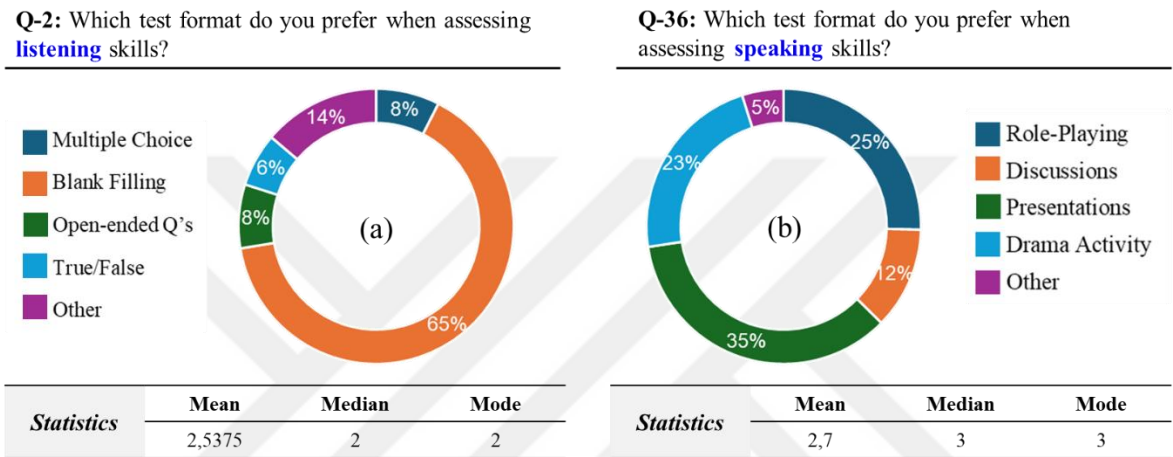
The statistical data presented in Figure 1 provide important insights into the frequency with which English teachers incorporate listening and speaking assessments into their examinations. As shown in Figure 1, the dominant practice among teachers is to examine these skills on a “*Per Semester*” basis. Of the respondents, 52.5% reported that they applied a listening assessment, and 51.2% noted that they performed a speaking assessment at least once per academic term. The results demonstrated that teachers recognize the significance of consistently evaluating oral competencies.

The “*monthly*” frequency of these assessments indicates a divergent trend. Of the teachers, 26.3% applied listening tests monthly and 31.3% performed speaking assessments monthly. This disparity suggests that teachers find it more practical to integrate speaking tests into standard evaluation practices than to integrate listening assessments. Conversely, “*Weekly*” evaluations of listening and speaking skills have rarely been conducted. Only 17% and 15% of the teachers reported conducting assessments at this frequency, respectively. The low rates may result from issues such as time constraints, curriculum requirements, or challenges associated with regularly assessing oral competencies. The percentage of teachers who perform “*Annual*” assessments compared to those who conduct no assessments is notably low (1.3%–2.5%).

Descriptive statistics further corroborate these findings. The mean scores for listening ($M = 2.45$) and speaking ($M = 2.4325$) evaluations were closely aligned, suggesting that teachers assessed these competencies with comparable regularity. The mode (3) and median (3) data further corroborate that the most often preferred assessment frequency is “*Per Semester*”.

Upon evaluating these findings collectively, it is evident that English teachers recognize the significance of measuring listening and speaking abilities; nevertheless, they do not perform these evaluations continuously, opting instead to conduct them at specific intervals. The diminished demand for weekly assessments highlights potential obstacles to consistent implementation. This circumstance may necessitate additional investigation to ascertain whether these evaluations may be conducted with greater frequency.

Figure 2: Statistics and Percentage Values for Q-2 & Q-36



The analysis of the preferred exam formats for assessing listening and speaking skills revealed distinct trends among the English teachers (Figure 2). In terms of listening assessments (Q-2), the vast majority of teachers (65%) prefer “*blank filling*” exercises. This indicates that teachers have a strong inclination towards structured question types that require students to recognize and process specific auditory information. Other formats, such as “*multiple-choice*” (8%), “*open-ended*” (8%), and “*true/false*” (6%), are much less preferred. This suggests that teachers perceive “*blank-filling*” exercises as a more effective or practical method for assessing listening skills.

On the other hand, speaking assessments (Item-36) exhibit a more varied distribution. The most preferred format is “*presentations*” (35%), followed by “*role-playing*” activities (25%) and “*discussions*” (23%). These results suggest that teachers prefer more interactive and communication-focused activities when evaluating their students' speaking skills. A possible reason behind this preference may be that such activities allow students to produce spontaneous and realistic language. In contrast, “*drama activities*” (12%) and other methods (5%) were used less frequently, indicating that teachers showed relatively less interest in more structured or scenario-based speaking tasks.

From a statistical perspective, the average value for listening assessments was 2.5375, with both the mode and median values being 2, which supports the idea that “*blank filling*” exercises are the most preferred format. In the speaking assessments, the average value was 2.7, with both the mode and median values being 3. This situation indicates that interactive and performance-based assessment formats (e.g., presentations and role-playing) are more widely adopted.

Overall, these findings indicate that structured and blank-filling-focused formats are given more importance in listening assessments, whereas interactive and communication-based tasks are prioritized in speaking assessments. This difference may reflect the nature of the skills being assessed because listening comprehension is generally evaluated with tests that measure specific points. In contrast, speaking assessments require more open-ended, dynamic interactions. These results highlight the importance of selecting formats that are suitable for the assessment objectives of each language skill.

Table 3: Statistics and Frequency Values for Q-3 & Q-37

Questions	Q3. Which factors do you think are important in assessing listening skills?		Q37. Which factors do you think are important in assessing speaking skills?	
Options	Frequency	%	Frequency	%
Task clarity	18	8,5	26	12,6
Material Quality	40	18,8	18	8,7
Time Management	46	21,6	42	20,3
Student Motivation	52	24,4	68	32,9
Measurement of Skills	55	25,8	53	25,6
Other	2	0,9	0	0
Total	213	100	207	100
Dichotomy group tabulated at value 1.				

The statistical data presented in Table 3 reveal the factors that English teachers considered important when evaluating listening (Q-3) and speaking (Q-37) skills. The results indicate significant differences in prioritization between the two skill areas.

In listening assessments (Q-3), the factor most frequently mentioned by teachers was “*Measurement of Skills*” (25.8%). This demonstrates that teachers prioritize the appropriate evaluation of students' understanding abilities through listening tests. “*Student Motivation*” (24.4%) and “*Time Management*” (21.6%) are also significant variables. The findings suggest that teachers evaluate the listening abilities of their students. Besides, they also recognize the importance of promoting “*Student Motivation*” and easing “*Time Management*”. “*Material Quality*” (18.8%)

and “Task Clarity” (8.5%) are considered less critical factors. Therefore, teachers’ perspectives are more acknowledged.

In the speaking evaluation (Q37), most of the teachers chose “*Student Motivation*” (32.9%) as the most important factor. This results in “*Student Motivation*” being the most effective factor in the speaking assessment. The speaking exams are interactive and performance-driven, so that the impact of the “*Student Motivation*” may be assumed to be a more crucial factor in assessing speaking. “*Measurement of Skills*” (25.6%) and “*Time Management*” (20.3%) are more significant factors, while “*Material Quality*” (8.7%) and “*Task Clarity*” (12.6%) are less important.

These data suggest that accurate measurement is the primary issue in the listening assessments. However, in the speaking assessments, the primary issue is “*Student Motivation*”. This divergence may occur as a result of the characteristics of speaking exams, including participation, self-confidence, and the readiness of students. Findings dictate that the focus should not solely be on skill measurement but also on the students’ reaction to the assessment procedure. Especially in the speaking assessment, increasing students’ motivation may lead to more effective and interactive exams.

The difficulties that teachers confront in their listening (Q-4) and speaking (Q-38) are given in Table 4. Considering the data in the table, both listening and speaking assessments share common challenges. However, some are not.

In the listening assessment (Q-4), teachers stated that the “*Medium Difficulty (MD)*” is the most common difficulty level. Especially, “*Time Management*” (43.75%), “*Material Compatibility*” (37.5%), and “*Student Involvement*” (40%) were stated as “*Medium Difficulty*”. In the listening assessment, these findings indicate that teachers struggle to manage time and prepare materials that cater to students’ varying levels. On the other hand, “*Assessment Objectivity*” (Mean = 2.4) was stated as a relatively “*Easy*”. Teachers are less concerned about “*Assessment Objectivity*” as they rated “*Very Easy*”.

Most of the teachers reported that in speaking assessments (Q-38), the “*Student Involvement*” (Mean = 3.48) is the greatest challenge. According to 33.75% of the teachers find “*Student Involvement*” as “*Medium Difficult*” while 26.25% of the teacher finds “*Students Involvement*” as “*Difficult*”. This indicates that student involvement in assessing speaking results poses a challenge to teachers. This challenge is due to student anxiety, lack of self-confidence, and classroom dynamics. Another challenging factor in speaking assessment is “*Material Compatibility*” (38.75% MD). The reported data in the table states that the selection of proper materials is also a problem.

Table 4: Statistics and Frequency Values for Q-4 & Q-38

Questions		Q4. Please indicate the degree of difficulty of the factors effective in assessing listening skills.					Q38. Please indicate the degree of difficulty of the factors effective in assessing speaking skills.				
Opt.	Level	Freq.	%	Mean	Med.	Mode	Freq.	%	Mean	Med.	Mode
Time Management	VE	4	5	3,15	3	3	11	13,75	3,3	3	3
	E	15	18,75				10	12,5			
	MD	35	43,75				22	27,5			
	D	17	21,25				18	22,5			
	VD	9	11,25				19	23,75			
	Total	80	100				80	100			
Material Compatibility	VE	6	7,5	3,013	3	3	13	16,25	2,75	3	3
	E	20	25				18	22,5			
	MD	30	37,5				31	38,75			
	D	15	18,75				12	15			
	VD	9	11,25				6	7,5			
	Total	80	100				80	100			
Student Involvement	VE	8	10	3,05	3	3	4	5	3,45	3	3
	E	16	20				11	13,75			
	MD	32	40				27	33,75			
	D	12	15				21	26,25			
	VD	12	15				17	21,25			
	Total	80	100				80	100			
Assessment Objectivity	VE	29	36,25	2,4	2	1	9	11,25	3,1	3	3
	E	13	16,25				17	21,25			
	MD	21	26,25				25	31,25			
	D	11	13,75				15	18,75			
	VD	6	7,5				14	17,5			
	Total	80	100				80	100			
VE: Very Easy (1), E: Easy (2), MD: Medium Difficulty (3), D: Difficult (4), VD: Very Difficult (5)											

Additionally, there was a significant difference between listening and speaking assessments in terms of objectivity. In listening assessments, objectivity was stated as easy to provide (mean = 2.4); however, in speaking assessments, it was much more challenging (mean = 3.1). The inherently subjective nature of speaking assessments may lead to different interpretations of components such as pronunciation, fluency, and coherence of expression by teachers.

In general, these findings suggest that both listening and speaking assessments present some challenges for teachers, but speaking assessments present more. In particular, ensuring student participation and making the assessment process more objective are of greater importance for

speaking examinations. Therefore, the reliability and fairness of speaking exams can be improved by developing well-structured assessment scales and guidelines with clear criteria.

Table 5: Statistics and Frequency Values for Q-5 & Q-39

Questions	Q5. Which factors do you think are important in assessing listening skills?		Q39. Which factors do you think are important in assessing speaking skills?	
Options	Frequency	%	Frequency	%
Student Performance	18	8,5	26	12,6
Exams	40	18,8	18	8,7
Projects	46	21,6	42	20,3
Portfolio Evaluation	52	24,4	68	32,9
Other	2	0,9	0	0
Total	213	100	207	100
Dichotomy group tabulated at value 1.				

The findings presented in Table 5 show which factors EFL teachers find most important while assessing listening (Q-5) and speaking (Q-39). The results present that “*Portfolio Evaluation*” methods are more preferred than traditional examinations.

In listening assessments (Q-5), “*Portfolio Assessment*” (24.4%) was found to be the most important factor. This means that it is more effective for teachers to monitor students’ performance continuously when assessing listening skills. “*Portfolio Assessment*” factor is followed by “*Projects*” (21,6%) and “*Exams*” (18,8%). These results suggest that traditional examination methods are still valid; however, there is a tendency towards alternative assessment practices. As an interesting fact, “*Student Performance*” (8.5%) had the lowest rate, and the teachers do not favor direct performance-based assessments.

A similar trend for speaking assessments (Q-39) as in listening assessments was observed. The “*Portfolio Assessment*” method resulted in a higher rate (32.9%) for the speaking assessment when compared to the listening one. This indicates that teachers attribute more importance to long-term performance. Monitoring. “*Projects*” (20.3%) and “*Student Performance*” (12.6%) are also highly considered. However, “*Exams*” (8.7%) had the lowest rate when assessing speaking skills.

In general, both listening and speaking assessments are replacing traditional examination methods with portfolio- and project-based assessment approaches. Generally, the reason teachers find these methods more effective is that they monitor students’ progress over time more accurately. These results are proof of the shift of today’s language assessments towards authentic

and process-oriented assessment methods. Teachers' tendency to this trend makes the student learning process more comprehensive and fairer.

Table 6: Statistics and Frequency Values for Q-7 & Q-41

Q7. Which of the following should be used to assess listening skills?			Q41. Which of the following should be used to assess speaking skills?		
Options	Frequency	%	Options	Frequency	%
Daily conversations/dialogs	78	33,3	Daily conversations/dialogs	78	32,8
Presentations	36	15,4	Presentations	57	23,9
Interviews	45	19,2	Interviews	35	14,7
Radio programs	33	14,1	Discussions	40	16,8
Music Tracks	42	17,9	Debates	28	11,8
Other	0	0	Other	0	0
Total	234	100	Total	238	100
Dichotomy group tabulated at value 1.					

The data about methods of listening (Q-7) and speaking (Q-41) assessment prepared by EFL teachers are presented in Table 6. Considering the results, EFL teachers prefer authentic and communication-based assessment techniques to structured and traditional assessment methods.

In listening assessments (Q-7), “*Daily conversations/dialogues*” (33.3%) was the most preferred technique. This shows that teachers think real-life interactive scenarios are more effective when assessing the listening skills of students. “*Interviews*” (19.2%) and “*Music Tracks*” (17.9%) were the other most preferred assessment methods. This indicates that a variety of auditory materials are important for assessing listening skills. On the contrary, “*Radio Programs*” (14.1%) and “*Presentations*” (15.4%) were less preferred. This means that teachers think that monologue-type auditory materials may be less effective in listening assessments.

In speaking assessments (Q-41), the most preferred method by teachers is “*Daily Conversations/Dialogues*” (32.8%). This indicates that EFL teachers think natural and spontaneous speaking is the best method when assessing students' speaking skills. “*Daily Conversations/Dialogues*” (32.8%). This is followed by “*Presentations*” (23.9%) and discussions (16.8%). This depicts that providing students with structured speaking opportunities is also important. On the contrary, “*Interviews*” (14.7%) and “*Debates*” (11.8%) were less preferred. This data suggests that formal and rule-based speaking tasks were found to be less effective by teachers.

Generally, these findings suggest that teachers prefer natural and interactive communicative methods in both listening and speaking assessments. This tendency is parallel with the CLT

approach. Teachers tend to use authentic assessment methods rather than structured tests to measure the language skills of students in real-life speaking scenarios.

When assessing language skills, besides accuracy and comprehension, it is necessary to take into consideration the natural use of language and communicative skills. The assessment methods preferred by teachers are approaches that support students in learning how to use language in the real world.

Table 7: Statistics and Frequency Values for Q-8 & Q-42

Questions	Q8. Which test format do you think is the most effective in assessing listening skills?		Q42. Which test format do you think is the most effective in assessing speaking skills?	
Options	Frequency	%	Frequency	%
Listening/Speaking and question-answer	13	16,3	44	55
Listening/Speaking and storytelling	4	5	7	8,75
Listening/Speaking and dialog completion	23	28,7	6	7,5
Listening/Speaking and comprehension questions	36	45	22	27,5
Other	4	5	1	1,25
Total	80	100	80	100
Statistics	Mean	3,175	Mean	2,113
	Median	3,5	Median	1
	Mode	4	Mode	1

The test formats preferred by EFL teachers in the assessment of listening (Q-8) and speaking (Q-42) skills are given in Table 7. Comprehension-oriented formats are preferred in listening assessments, while interaction-based formats are preferred in speaking assessments.

In listening assessments (Q-8), the most preferred test format by teachers was “*Listening and Comprehension Questions*” (45%). This means that teachers focus on students' ability to process and interpret auditory inputs. This format was followed by “*Dialogue Completion*” (28.7%). This suggests that teachers also consider students' ability to make sense of the context when assessing listening skills. On the contrary, formats such as “*Question-Answer*” (16.3%) and “*Storytelling*” (5%) were not frequently preferred by teachers in listening assessments.

In speaking assessments (Q-42), the “*Question-Answer*” (55%) test format was far the most preferred by the EFL teachers. This means that teachers find speaking tasks based on direct interaction more effective. The second most preferred method was “*Comprehension Questions*”

(27.5%). This indicates that teachers wanted to assess both listening and speaking skills of students together. Conversely, the “*Storytelling*” (8.75%) and “*Dialogue Completion*” (7.5%) formats were less preferred. “*Storytelling*” and “*Dialogue Completion*” are found to be less practical or reliable methods by EFL teachers.

The mean score of listening assessments (mean value = 3.175) was higher than that of speaking assessments (mean value = 2.113). This is because teachers accept listening assessments as a process that requires processing and interpretation. Furthermore, the most frequently selected test format for listening assessments (mode value = 4) was higher than that of speaking assessments (mode value = 1). This means confirming that meaning-oriented approaches were preferred for listening assessments, whereas direct interaction-oriented approaches were preferred for speaking assessments.

Results suggest that teachers prefer comprehension-based listening assessments, while their preference is for interaction-based formats in speaking assessments. This also emphasizes the importance of a balanced assessment of students by covering both comprehension and performance.

In Table 8, responses of EFL teachers on the strategies used in listening (Q-9) and speaking (Q-43) skills assessments. According to the results, the variety of test types is more important in listening, while the variety of content and tasks is more important in speaking.

In listening assessments (Q-9), the teachers most frequently prefer “*Use a variety of exam formats*” (43.8%) as an assessment strategy. This finding leads teachers to believe that applying different types of tests can increase the reliability of listening assessments. It is thought that assessing students with various question types and configurations provides a more comprehensive listening skill assessment. In addition, the strategy of “*Providing student feedback regularly*” (21.3%) is another listening assessment strategy. This means that teachers consider feedback as critical for the development of listening skills. On the other hand, strategies such as “*Using Sample Materials*” (18.8%) and “*Providing content and task variety*” (15%) were less preferred.

In speaking assessments (Q-43), teachers believe that “*Providing content and task variety*” (31.3%) is the most important strategy. This finding leads us to the most effective strategy for assessment of speaking skills. These are assigning a variety of tasks involving different communication situations and types of interactions. The high preference rate of “*Providing Regular Feedback to Students*” (26.3%) strategy means that feedback is considered a critical skill in speaking development. Interestingly, strategies such as “*Using sample materials*” (22.5%) and “*Use a variety of exam formats*” (18.8%) were preferred at lower rates, suggesting that teachers focus more on task content rather than format variety in speaking skill assessments.

Table 8: Statistics and Frequency Values for Q-9 & Q-43

Questions	Q9. Which strategies do you think are more effective in your listening tests?		Q43. Which strategies do you think are more effective in your speaking tests?	
Options	Frequency	%	Frequency	%
Using sample material	15	18,8	18	22,5
Use a variety of exam formats	35	43,8	15	18,8
Providing Content and Task Variety	12	15	25	31,3
Providing student feedback regularly	17	21,3	21	26,3
Other	1	1,3	1	1,3
Total	80	100	80	100
Statistics	Mean	2,425	Mean	2,65
	Median	2	Median	3
	Mode	2	Mode	3

From a statistical point of view, the mean score for speaking assessment strategies (Mean Value = 2.65) was slightly higher than that for listening assessment strategies (Mean Value = 2.425). This finding supports the notion that speaking assessments require more contextual variability. Furthermore, the median and mode values confirmed that teachers preferred more interactive and content-enriched methods in speaking assessments.

In general, these findings show that teachers determine their assessment strategies according to the nature of the assessed skills. The diversity of test formats in listening assessments is seen as a factor that increases reliability. In contrast, the diversity of content and tasks in speaking assessments comes to the fore, revealing that teachers adopt different approaches to assess both skills effectively.

The statistical data presented in Table 9 reveal EFL teachers' perceptions of scoring consistency in the listening (Q-10) and speaking (Q-44) assessments. The results show that teachers have a particular perception of reliability in the assessment processes for both skills. However, they think that speaking assessments contain more variability than listening assessments do.

In listening assessments (Q-10), the majority of the teachers (67.5%) stated that they believed that if another teacher scored the same test, *“he/she would give a similar score”*. In addition, 27.5% thought that *“He/she would give the same score”*. Only 3.8 per cent indicated that another teacher would give a lower score, indicating that listening assessments generally had a high level of consistency. This can be explained by the fact that listening tests usually have a more structured format that allows for more objective scoring.

Table 9: Statistics and Frequency Values for Q-10 & Q-44

Questions	Q10. Do you think another teacher would have given similar or the same points if he/she had assessed your listening test?		Q44. Do you think another teacher would have given similar or the same points if he/she had assessed your speaking test?	
Options	Frequency	%	Frequency	%
He/she would give a higher score.	1	1,3	3	3,8
He/she would give a similar score.	54	67,5	58	72,5
He/she would give the same score.	22	27,5	5	6,3
He/she would give a lower score.	3	3,8	14	17,5
Total	80	100	80	100
Statistics	Mean	2,3375	Mean	2,375
	Median	2	Median	2
	Mode	2	Mode	2

A significant tendency was noted in the speaking assessments (Q-44). While 72.5% of respondents indicated that another teacher would provide an equivalent score (*He/she would give a similar score*), only 6.3% claimed that another teacher would assign the same score (*He/she would give the same score*), indicating a greater variety in speaking assessments. Furthermore, 17.5% indicated that yet another teacher would give a lower score (*He/she would give a lower score*), significantly higher than the 3.8% observed for listening evaluations. This can be explained by the inherent subjectivity of speech assessments, whereby teachers assign varying levels of significance to different variables. The mean results for both listening ($M = 2.3375$) and speaking ($M = 2.375$) evaluations were closely aligned, suggesting that teachers typically trust the reliability of the scoring. The increased variety in speech assessments indicates that subjective factors significantly affect the assessment procedures.

These findings suggest that speaking assessments contain more scoring inconsistencies than listening assessments do, which may be due to differences in interpretation among teachers. The fact that speaking skills consist of different components, such as fluency, pronunciation, grammatical accuracy, and completeness of expression, may cause scoring to be subjective. On the other hand, the fact that listening assessments include more structured question types may facilitate objective assessments and increase scoring consistency. Results emphasize the need for more standardized scoring criteria and clear rubrics in speaking assessments. Organizing training to increase scoring consistency among teachers and standardizing assessment processes can contribute to making speaking exams more reliable.

Table 10: Statistics and Frequency Values for Q-11 & Q-45

Questions	Q11. To what extent do you think the results of your listening test predict real-world listening performance?		Q45. To what extent do you think the results of your listening test predict real-world speaking performance?	
Options	Frequency	%	Frequency	%
Predicts at a very high level.	4	5	4	5
Predicts at a high level.	16	20	17	21,3
Predicts at a medium level.	41	52,2	41	51,2
Predicts at a low level.	18	22,5	15	18,8
Does not predict at all.	1	1,3	3	3,6
Total	80	100	80	100
Statistics	Mean	2,95	Mean	2,95
	Median	3	Median	3
	Mode	3	Mode	3

The statistical data presented in Table 10 reveal EFL teachers' perceptions of the extent to which listening (Q-11) and speaking (Q-45) assessments reflect students' real-world language performances. The results indicate that these assessments are believed to reflect students' communication skills to some extent but are generally considered to have a moderate level of validity.

In the listening assessments (Q-11), the majority of teachers (52.2%) indicated that they thought that the listening exams reflected students' real-world listening skills to a moderate extent (*Predicts at a medium level*). However, 20% stated that these assessments had a high level of reflectivity (*Predicts at a high level*), and 5% stated that they had a very high level of reflectivity (*Predicts at a very high level*). On the other hand, 22.5% thought that listening assessments had a low level of reflectivity (*Predicts at a low level*), while 1.3% stated that they did not reflect at all (*Does not predict at all*). These findings suggest that listening tests may not fully reflect students' real-life listening skills.

A similar trend was observed in the speaking assessments (Q-45). 51.2% of the teachers believed that speaking exams reflected students' real-world speaking skills at a moderate level, "*Predicts at a medium level*". On the other hand, 21.3% of the teachers stated that the assessments had a high level of reflectivity "*Predicts at a high level*", and 7.5% stated that they reflected at a very high level, "*Predicts at a very high level*". However, 18.8% felt that the speaking assessments had low reflectivity, "*Predicts at a low level*", and 3.6% felt that they did not reflect real-world performance at all (*Does not predict at all*).

Statistically speaking, the mean scores of the listening and speaking assessments ($M = 2.95$) were quite close to each other, indicating that the assessment processes for both skills had a similar perception of validity. Furthermore, the median and mode values (both three) confirm that the most common perception is ‘moderate reflectivity.’

These findings suggest that teachers do not have absolute confidence that listening and speaking assessments accurately reflect students' real-world language skills. Especially in speaking assessments, due to the nature of the test environment, teachers may be unable to measure student performance in real communication situations fully. Similarly, the fact that the materials and question types used in listening assessments do not fully match the auditory inputs that students encounter in real life may contribute to these limitations.

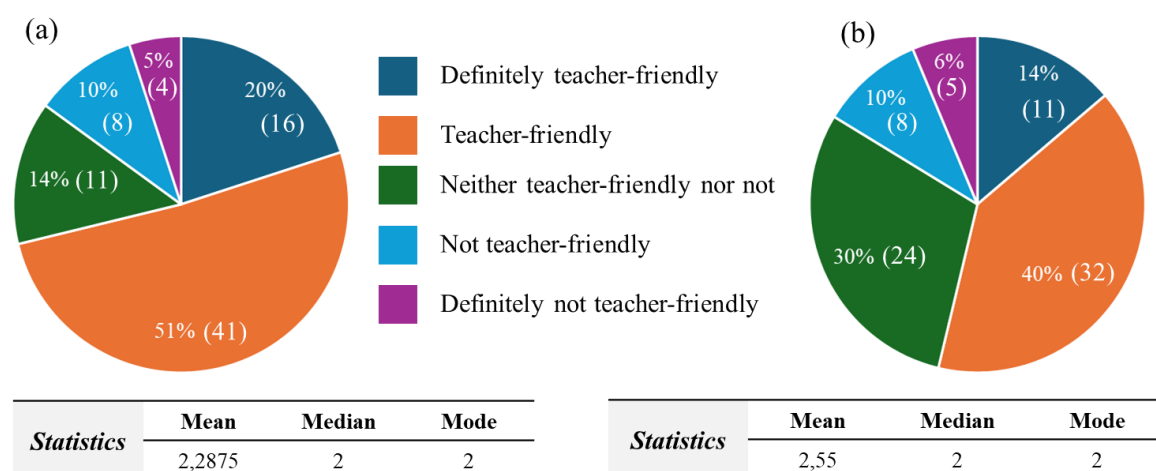
These results indicate that language tests should be more realistic and closely tied to real-life situations. Using tasks and scenarios similar to the language that students use in daily life experiences may be more valid.

The statistical data on the perception of EFL teachers regarding the feasibility of listening (Q-12) and speaking (Q-46) assessments, as well as the ease of the assessment process, are presented in Figure 3. The data show that both types of examinations are generally considered feasible by teachers. However, teachers believe that listening examinations are easier to manage than speaking examinations.

Figure 3: Statistics and Frequency Values for Q-12 & Q-46

Item-12: In your opinion, were the **listening** tests you administered generally teacher-friendly (easy to administer and evaluate)?

Item-46: In your opinion, were the **speaking** tests you administered generally teacher-friendly (easy to administer and evaluate)?



In the listening assessments (Q-12), 51% of the teachers rated these exams as “*Teacher-friendly*.” In addition, 20% rated the listening tests as “*Definitely teacher-friendly*,” indicating that the administration and assessment of listening tests are generally practical for teachers. However, 14% of the teachers stated that the listening tests were “*Not teacher-friendly*,” suggesting that some teachers experienced difficulties in their listening assessments, such as technical problems, time constraints, or lack of appropriate aural materials.

In speaking exams (Q-46), merely 40 percent of teachers characterized them as “*Teacher-friendly*,” a figure that is lower than that of listening assessments. Moreover, only 14 percent classified speaking tests as “*Definitely teacher-friendly*,” suggesting that teachers perceive the administration and evaluation of speech evaluations as more complex. Conversely, 30% of the teachers classified the speaking tests as “*Neither teacher-friendly nor unfriendly*,” indicating a variability in teachers' experiences about the applicability of these assessments. The proportion of individuals who considered speaking assessments as “*Not teacher-friendly*” exceeded that of those who regarded listening tests as “*Not teacher-friendly*.”

The mean score for speaking examinations ($M = 2.55$) exceeded that for listening examinations ($M = 2.2875$), indicating that speaking exams pose more management challenges for teachers. Moreover, the median and mode scores for both exam types were identical (2), suggesting that most teachers deemed both exams broadly relevant, although speaking exams presented a marginally greater difficulty.

Overall, these findings suggest that EFL teachers find listening assessments to be more manageable than speaking assessments. It is thought that the greater difficulty in administering and assessing speaking examinations may be due to the greater role of subjective elements in scoring, the time-consuming nature of the assessment process, and the logistical complexity of organizing oral examinations for large groups of students.

Therefore, methods such as the use of standardized assessment rubrics, automated scoring tools, and the development of alternative assessment formats should be considered to make speaking assessment processes more feasible for teachers.

The statistical data presented in Figure 3 show the perception of EFL teachers regarding the feasibility of listening (Q-12) and speaking (Q-46) assessments, as well as the ease of the assessment process reveal EFL teachers' perceptions of the reliability (consistency) and validity (the degree to which they measure the skills they are supposed to measure of listening (Q-13 & Q-47) and speaking (Q-14 & Q-48) assessments. The findings show that listening assessments were generally perceived to be more reliable, but there was more uncertainty about the validity of speaking assessments.

Table 11: Statistics and Frequency Values for Q-13, Q-14, Q-47 & Q-48

Questions	Q13. Do you think that your listening tests give sufficiently reliable (consistent) results?		Q14. Were the speaking tests you conducted sufficiently valid (did they measure what was intended)?		Q47. Do you think that your listening tests give sufficiently reliable (consistent) results?		Q48. Were the speaking tests you conducted sufficiently valid (did they measure what was intended)?	
Options	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Definitely yes	18	22,5	4	5	16	20	5	6,3
Yes, mostly	39	48,8	42	52,5	41	51,2	42	52,5
Not sure	13	16,3	25	31,3	11	13,8	21	26,3
No, not really	5	6,3	6	7,5	8	10	8	10
Absolutely not	5	6,3	3	3,8	4	5	4	5
Total	80	100	80	100	80	100	80	100
Statistics	Mean	2,25	Mean	2,525	Mean	2,2875	Mean	2,55
	Med.	2	Med.	2	Med.	2	Med.	2
	Mode	2	Mode	2	Mode	2	Mode	2

The statistical data presented in Figure 3 show the perception of EFL teachers regarding the feasibility of listening (Q-12) and speaking (Q-46) assessments, as well as the ease of the assessment process reveal EFL teachers' perceptions of the reliability (consistency) and validity (the degree to which they measure the skills they are supposed to measure) of listening (Q-13 and Q-47) and speaking (Q-14 and Q-48). The findings indicate that listening assessments are generally perceived as more reliable, but there is greater uncertainty regarding the validity of speaking assessments.

For the listening assessments (Q-13 and Q-47), about half of the teachers (48.8% and 51.2%) reported that these tests were “*mostly reliable*,” indicating a general confidence that listening tests are generally administered stably and consistently. In addition, 22.5% and 20% rated the listening tests as “*definitely reliable*” respectively, suggesting that teachers have a high belief in the reliability of listening assessments. However, 16.3% were unsure, and 6.3%-8.8% did not find the listening tests reliable, suggesting that some teachers may have doubts about the consistency of the listening tests.

A different trend was observed in the speaking assessments (Q-14 and Q-48). 52.5% of the teachers think that speaking exams are “*mostly valid*,” but 31.3% of the teachers are undecided (*not sure*). This is considerably higher than the rate of uncertainty observed in listening assessments, suggesting more uncertainty among teachers regarding the validity of speaking tests. In addition, 10% thought that speaking tests were “*not very valid*.” In comparison, 6.3% thought that they were

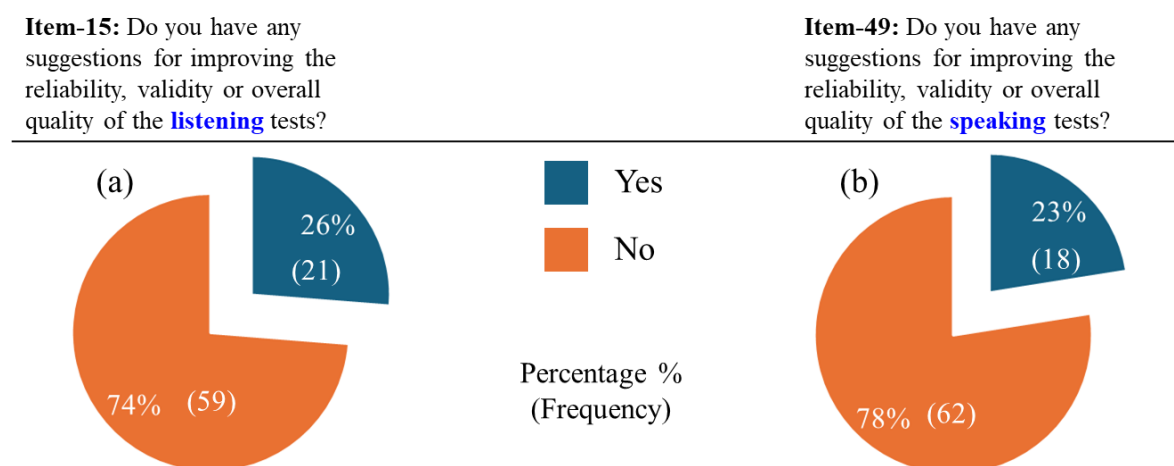
“*definitely not valid*,” suggesting that speaking assessments are more questioned in terms of validity than reliability.

Statistically speaking, the reliability means of the listening tests ($M = 2.25$ and $M = 2.2875$) were slightly lower than the validity means of the speaking tests ($M = 2.525$, $M = 2.55$). This suggests that teachers were more concerned about the validity of the speaking tests but had more confidence in the reliability of the listening tests. In addition, the mode values of 2 for both skills indicate that teachers are generally concentrated in the “mostly reliable” or “mostly valid” categories.

These findings suggest that, in general, teachers believe that listening assessments are more reliable but that validating speaking assessments is a more complex and uncertain process. The skills assessed in speaking tests are inherently open to subjective interpretation, which may lead to differences in assessments among teachers. However, the fact that listening tests usually contain structured questions may make scoring more objective and reliable. Results suggest that to make speaking assessments more reliable and valid, standardized rubrics should be developed, teacher training should be increased, and assessment processes should be made more explicit.

The statistical findings presented in Figure 4 reveal that EFL teachers' perceptions of listening (Q-15) and speaking (Q-49) assessments need to be improved in terms of reliability, validity, and overall quality. The results show that most teachers find the current assessment methods adequate, but a small group of them think that improvements need to be made.

Figure 4: Statistics and Frequency Values for Q-15 & Q-49



A similar trend was observed in the speaking assessments (Q-49). 78% of the teachers stated that no changes were needed in speaking assessments. However, 23% argued that the speaking assessment processes should be improved. Although this percentage is slightly lower than the

suggestions for improvement made for listening assessments, it indicates that, overall, teachers largely accept speaking exams in their current form.

Statistically, a slightly higher proportion of teachers believe that listening tests need improvement (26%) compared to speaking tests (23%). This slight difference suggests that teachers more frequently question certain aspects of listening tests; however, both types of assessments are generally regarded as adequate overall. These results lead to general satisfaction with current assessment methods. However, a specific group of teachers feels that the examination processes should be improved.

Generally, these findings show that most teachers find listening and speaking assessments effective in their current form. However, it is crucial to review and update evaluation methods regularly. Focusing on the areas teachers see as needing improvement helps make assessment approaches more fair, supportive of students, and better for developing practical skills.

The statistical findings presented in Figure 5 reveal EFL teachers' overall satisfaction with practice tests measuring listening (Q-17) and speaking (Q-51) skills. The findings show that teachers with slightly higher satisfaction rates evaluate listening tests, but there is more uncertainty regarding speaking tests. In the listening assessments (Q-17), 34% of the teachers indicated that they were "*somewhat satisfied*" with the tests. Additionally, 21% expressed that they were "*very satisfied*" with the listening tests, suggesting that teachers generally perceived listening evaluation methods to be largely effective. Nonetheless, 10% of participants reported being "*not very satisfied*," and another 10% indicated they were "*not at all satisfied*," implying that some teachers might have particular issues regarding the structure, material, or evaluation standards of the listening assessments.

In the speaking assessment (Q-51), the degree of satisfaction exhibited a somewhat different pattern. A broader range of opinions regarding speaking exams was shown by the most frequently given response, "*neither satisfied nor dissatisfied*" (38%). Teachers were "*somewhat satisfied*" with the speaking evaluations, according to 36% of them, while the percentage of "*very satisfied*" teachers (16%) was lower than that of the listening assessments. Regarding the speaking tests, all of the teachers opted not to choose the "*not at all satisfied*" option, indicating that teachers generally approve of these tests. This variation suggests a more consistent perception of satisfaction with the listening tests, whereas teachers' varying experiences and assessments influence the speaking tests. Additionally, the variation in mode values (mode = 2 for listening assessments and mode = 3 for speaking assessments) indicated that there were more neutral or indecisive views regarding speaking assessments.

Figure 5: Statistics and Frequency Values for Q-17 & Q-51

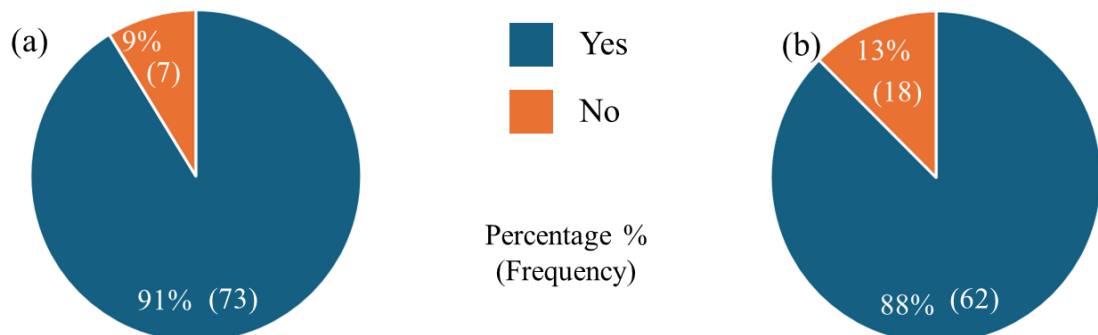
	Item-17: What is your overall satisfaction with the listening skills practice test?		Item-51: What is your overall satisfaction with the speaking skills practice test?	
Options	Percentage	Frequency	Percentage	Frequency
<i>Very satisfied</i>	21%	17	16%	13
<i>Somewhat satisfied</i>	34%	27	36%	29
<i>Neither satisfied nor dissatisfied</i>	25%	20	38%	30
<i>Not very satisfied</i>	10%	8	10%	8
<i>Not satisfied</i>	10%	8	0%	0
Statistics	Mean:	2.5375	Mean:	2.4125
	Median:	2	Median:	2
	Mode:	2	Mode:	3

In general, these findings suggest that teachers find both types of assessments acceptable, but there is more uncertainty regarding speaking tests. Teachers' satisfaction levels could be improved by establishing more standardized criteria for the assessment and management of speaking tests. Especially, well-defined assessment criteria, student-specific feedback, and structured scoring rubrics for speaking assessments may help teachers find speaking tests more reliable and effective.

Figure 6: Statistics and Frequency Values for Q-18 & Q-52

Item-18: Did you provide support to students using the **listening** test results?

Item-52: Did you provide support to students using the **speaking** test results?



In Figure 6, the question is whether EFL teachers support students using listening (Q-18) and speaking (Q-52) test results. The rate values in the figure depict that the majority of teachers integrate assessment results into their teaching processes and provide feedback to students.

In the listening assessments (Q-18) data, 91% of the teachers stated that they provided support to their students by using the listening test results. This high rate shows that teachers use listening tests not only in assessing students, but also as feedback data to improve their listening skills. On the other hand, a small group of teachers (9%) stated that they did not use the listening exam results as feedback data to students.

As for speaking assessments (Q-52), the proportion of teachers (88%) uses speaking test results to support students' speaking skills. On the contrary, 13% of the teachers stated that they did not use speaking test results to support students' speaking skills. This suggests that teachers may have difficulty providing support from the test results of speaking assessments when compared to listening assessments. This may be the result of grading being much easier in a listening assessment than in a speaking one.

From a statistical point of view, both listening and speaking test results are used to provide support to students by teachers. However, this rate is slightly lower for speaking assessments (91% vs. 88%). This slight reason may be because of difficulties in providing feedback on speaking exams. Examining these data, it is evident that teachers use test results not only to determine students' success in exams and grade them, but also to help them improve their language skills.

Table 12 presents the types of support that English teachers provide to students based on their listening (Q-19) and speaking (Q-53) test results. Upon reviewing the table, teachers employ various methods to enhance listening and speaking skills. There are significant differences in the support strategies for both skills.

Table 12: Statistics and Frequency Values for Q-19 & Q-53

Questions	Q19. What kind of support did you provide to the students using the listening test results?		Q53. What kind of support did you provide to the students using the speaking test results?	
	Frequency	%	Frequency	%
Additional course materials	28	22,8	24	34,3
Individual counselling	28	22,8	38	54,3
Group work	17	13,8	34	48,6
Feedback sessions	50	40,7	37	52,9
Other	0	0	0	0
Total	123	100	133	100
Dichotomy group tabulated at value 1.				

"Feedback sessions" (40.7%) were the most frequently used support strategy in listening assessments (Q-19). This rate suggests that teachers aim to improve students' listening skills. Therefore, teachers would rather provide them with structured, direct feedback. Additionally,

teachers provide support for the listening skills of students by offering them “*Additional course materials*” (22.8%) and “*Individual counselling*” (22.8%). Nevertheless, the least popular approach was “*Group work*” (13.8%), indicating that listening abilities are typically evaluated according to personal growth.

In speaking assessments (Q-53), the trend differs from that in the listening assessments. The “*Individual Counselling*” option was chosen mainly by teachers (54.3%). This suggested that one-on-one teaching feedback is necessary to improve speaking abilities. Furthermore, “*Feedback sessions*” (52.9%) and “*Group work*” (48.6%) were strongly favored. This means that effective communication and teamwork are essential for enhancing speaking abilities. On the other hand, the method of “*Additional course materials*” (34.3%) was preferred at a lower rate compared to listening assessments. That is why the development of speaking skills requires more direct interaction.

From a statistical point of view, it is seen that support strategies for listening test results are primarily based on structured feedback and additional resources. At the same time, interaction-based methods, such as individual guidance and group work, are used more in speaking tests. This suggests that teachers find individual-oriented and material-supported methods more effective for developing listening skills, whereas interactive and personalized approaches are more effective for developing speaking skills.

All in all, these findings reveal that teachers adopt different support approaches for both skills. While assessments of listening skills are generally based on individual work and structured feedback processes, more interactive and personalized support strategies are preferred for developing speaking skills. These results suggest that the most effective support mechanisms for both skills should be identified, and training programs should be organized for teachers to make these processes more efficient.

The statistical findings presented in Figure 7 reveal that EFL teachers' perceptions of the extent to which listening (Q-20) and speaking (Q-54) assessments contribute to improving student motivation. The findings indicate that both skill assessments increase students' motivation to a certain extent, but speaking assessments have a more substantial effect.

For the listening assessments (Q-20), 43% of teachers think that these exams motivate students to improve their listening skills “*moderately motivated*”. In addition, 24% stated that listening assessments were “*highly motivating*” for students. However, only 9% found them “*very highly motivating*.” On the other hand, 18% said that listening assessments were low motivating, and 8% said that they had no impact on students' motivation. This suggests that listening assessments have a specific effect on motivation, although it is not particularly strong.

Figure 7: Statistics and Frequency Values for Q-20 & Q-54

	Item-20: To what extent did the listening tests motivate students to improve their skills?		Item-54: To what extent did the speaking tests motivate students to improve their skills?	
Options	Percentage	Frequency	Percentage	Frequency
<i>Very highly motivated</i>	9%	7	10%	8
<i>Highly motivated</i>	24%	19	33%	26
<i>Moderately motivated</i>	43%	34	51%	41
<i>Motivated at a low level</i>	18%	14	0%	0
<i>Not motivated at all</i>	8%	6	6%	5
Statistics	Mean:	2.9125	Mean:	2.6
	Median:	3	Median:	3
	Mode:	3	Mode:	3

In speaking assessments (Q-54), the students' motivation levels were evaluated as higher. While 51% of the teachers stated that speaking tests motivated students at a “*moderate level*,” 33% of them considered speaking tests as “*highly motivating*.” The percentage of those who found them highly motivating (10%) was slightly higher than that of the listening assessments. One of the most striking differences is the comparison of the proportions rated as “low motivating.” While this rate was 18% in the listening assessments, no teacher in the speaking assessments considered speaking tests to be a low-motivating method.

Statistically speaking, the average motivation score for the listening tests ($M = 2.9125$) was slightly lower than that for the speaking tests ($M = 2.6$). This difference suggests that speaking assessment has a more substantial impact on student motivation. However, the median and mode values for both types of exams were the same (3), indicating that teachers generally regarded these assessments as “*moderately motivating*.”

Overall, these results imply that, while speaking assessments are thought to be a more powerful source of incentive, listening assessments also appear to have some effect on student motivation. The fact that speaking exams by their very nature demand active engagement from students and offer direct communication practice may help to explain this discrepancy. However, the impact of listening exams on motivation may be limited because they focus on a more passive assessment of skills, and do not directly promote students' productive engagement.

These findings suggest that to boost motivation, listening assessments, in particular, need to be made more interactive. It might be helpful to make listening assessments more authentic, centered on student engagement, and incorporating relevant situations, to guarantee that students are motivated while developing their listening abilities. However, assessment styles and feedback

methods that help students feel more at ease can be used to enhance the motivational effect of speaking exams further.

Figure 8: Statistics and Frequency Values for Q-21 & Q-55

	Item-21: How much do you think listening tests affect teaching and learning?		Item 55: How much do you think speaking tests affect teaching and learning?	
Options	Percentage	Frequency	Percentage	Frequency
<i>Very positive</i>	9%	7	9%	7
<i>Positive</i>	49%	39	43%	34
<i>Neither positive nor negative</i>	34%	27	33%	26
<i>Negatively affects</i>	8%	6	15%	12
<i>Very negatively affects</i>	1%	1	1%	1
Statistics	Mean:	2.4375	Mean:	2.5750
	Median:	2	Median:	2
	Mode:	2	Mode:	2

The statistical findings presented in Figure 8 reveal EFL teachers' perceptions of the impact of listening (Q-21) and speaking (Q-55) tests on their teaching and learning process. The results show that both types of examinations have an overall positive impact; however, speaking assessments cause more difficulties and mixed opinions.

In the case of listening assessments (Q-21), 49% of teachers indicated that these exams had a “*positive*” impact on the teaching and learning processes. Furthermore, 9% of the participants thought that this effect was “*very positive*”, suggesting that listening assessments are typically regarded as a kind of assessment that aids in language acquisition. Nonetheless, 34% of teachers assessed the influence of hearing assessments as “*neither positive nor negative*”, suggesting that some teachers are unsure of the function of listening assessments in the teaching and learning processes. However, according to 8% of the respondents, listening tests had a “*negative*” impact on instruction and learning, and 1% said that their influence was “*very negative*”.

A similar but mixed trend was observed in speaking assessments (Q-55). 43% of the teachers think that speaking tests have a “*positive*” effect on the teaching process, and 9% think that they have a “*very positive*” effect. However, the percentage of those who found speaking tests “*negative*” (15%) was higher than that of listening assessments. In addition, the proportion of those (33%) who considered the impact of speaking tests as “*neither positive nor negative*” was quite similar to that of listening tests.

In terms of statistics, the speaking tests' average teaching impact score ($M = 2.5750$) was marginally greater than that of the listening tests ($M = 2.4375$). This discrepancy suggests that

speaking tests have a significant impact on the educational process, although this influence can be interpreted in both favorable and unfavorable ways. Teachers viewed these tests as “*positive*” in general, as evidenced by the fact that the mode and median values for both exam kinds were 2.

In general, these findings reveal that listening and speaking exams contribute to the teaching and learning processes, but speaking exams cause more controversial opinions. It can be thought that speaking exams may put more pressure on students due to their nature, and subjective factors may be effective in the evaluation process. However, a more structured and predictable format of listening tests may offer a more stable assessment process for teachers.

These results suggest that to increase the contribution of speaking assessments to the teaching process, more structured feedback mechanisms, student-centered assessment techniques, and test formats that provide low-anxiety environments should be developed. In particular, alternative assessment methods (e.g., collaborative speaking activities or project-based assessments) can be implemented to make students feel more comfortable.

Figure 9: Statistics and Frequency Values for Q-22 & Q-56

	Item-22: Do you think that the listening tests you conduct contribute sufficiently to students' language development?		Item 56: Do you think that the speaking tests you conduct contribute sufficiently to students' language development?	
Options	Percentage	Frequency	Percentage	Frequency
Definitely, Yes	19%	15	16%	13
Yes, mostly	43%	34	39%	31
Not sure	24%	19	26%	21
No, not really	11%	9	16%	13
Absolutely not	4%	3	3%	2
Statistics	Mean:	2.3875	Mean:	2.5000
	Median:	2	Median:	2
	Mode:	2	Mode:	2

The statistical findings presented in Figure 9 reveal English teachers' perceptions of the extent to which listening (Q-22) and speaking (Q-56) exams contribute to their students' language development. The results indicate that both types of exams generally contribute positively to language development, but speaking assessments tend to elicit more varied opinions.

According to 43% of teachers, listening tests had a “*Yes, mostly*” effect on children’s language development (Q-22). Furthermore, according to 19% of the respondents, listening tests have a “*Definitely, yes*” influence, suggesting that most teachers view them as tools that aid language development. However, according to 24% of the teachers, the impact of listening tests is

“*Not sure*,” suggesting that some teachers think the tests have little to no impact on language development. However, according to 11% of the respondents, speaking tests had a “*No, not really*” influence, while 4% said that they had a “*Absolutely not*” effect.

In speaking assessments (Q-56), however, a similar but more variable trend was observed. 39% of the teachers believe that speaking exams have a “*Yes, mostly*” effect on language development, while 16% think they provide a “*Definitely yes*” contribution. However, the percentage of those who find speaking exams “*No, not really*” (16%) is higher than that of listening assessments. Additionally, the percentage of those who evaluated the impact of speaking exams as ‘*Not sure*’ (26%) was slightly higher than that of listening exams. This suggests that teachers have varying experiences regarding the impact of speaking exams.

In terms of statistics, speaking tests had a slightly higher average score ($M = 2.5$) than listening tests ($M = 2.3875$), in terms of their influence on language development. This discrepancy shows that speaking tests have a greater influence on language development, but also raises the possibility that some teachers may not see this effect favorably. Both exam types’ median and mode values of 2 attest to the fact that teachers generally consider these evaluations to be “*Yes mostly*.”

All things considered, these results indicate that speaking tests elicit more divergent viewpoints, while listening tests also contribute to language development. This discrepancy is believed to be caused by increased anxiety that speaking tests can induce in children, disparities in teacher grading, and larger individual variations in speaking abilities across students.

These findings suggest that to improve speaking evaluations, more organized feedback systems, student-centered evaluation methodologies, and techniques for establishing real-world speaking situations are required. Alternative techniques, such as group speaking exercises, peer evaluations, and project-based assessments, can be incorporated into the assessment procedure to motivate students to improve their speaking abilities.

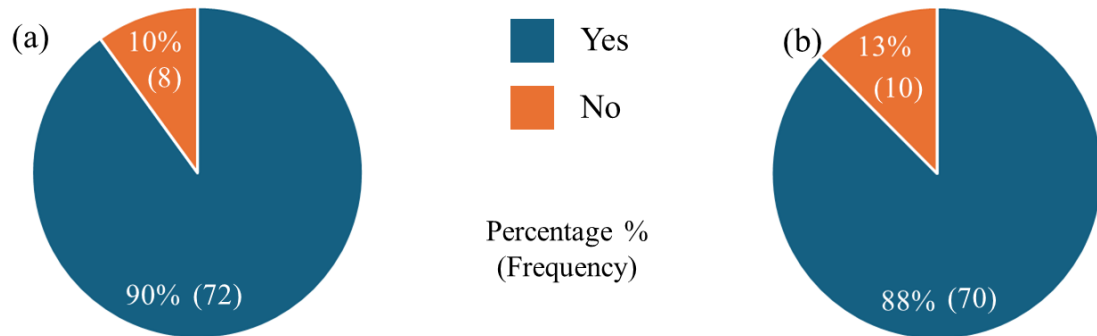
The data in Figure 10 presents how much English teachers include student feedback in listening (Q-23) and speaking (Q-57) assessments. The results show that most teachers consider student feedback important and use it more frequently in listening exams.

Ninety percent of the teachers reported that they consider student feedback in their listening assessments (Q-23). This high percentage suggests that teachers find it fairer and more effective to integrate student feedback into listening exams. The majority of teachers appreciate student feedback in listening assessments. Ten percent of teachers, who are not integrating student feedback into listening evaluation, prove that.

Figure 10: Statistics and Frequency Values for Q-23 & Q-57

Item-23: Do you use student feedback in the **listening** test evaluation process?

Item-57: Do you use student feedback in the **speaking** test evaluation process?



In speaking assessments (Q-57), however, teachers slightly less integrate the student feedback into the evaluation process (88%). On the other hand, 13% of teachers do not consider student feedback while administering speaking evaluations. This difference may be because the addition of feedback to speaking tests is more difficult. Integrating student feedback into the grading process can be challenging, as speaking tests are inherently subjective.

According to statistics, teachers typically integrate student feedback into the evaluation process for speaking and listening tests. This trend is slightly considered for speaking exams. This difference might be explained by the fact that feedback can be assessed using objective standards. Listening tests are more systematic, while listening speaking exams are more subjective. Although less typical in speaking tests, these data generally show that teachers are eager to integrate student feedback into the assessment process.

Table 13: Statistics and Frequency Values for Q-24 & Q-58

Questions	Q24. How do you use student feedback in the listening exam assessment process?		Q58. How do you use student feedback in the speaking exam assessment process?	
Options	Frequency	%	Frequency	%
By providing suggestions for improvement	43	30,1	45	31,5
Emphasizing students' strengths and weaknesses	53	37,1	46	32,2
By preparing progress reports	10	7	15	10,5
By conducting student-teacher interviews	37	25,9	37	26,9
Other	0	0	0	0
Total	80	100	143	100

In Table 13, the trends in how English teachers employ student feedback in their listening (Q-24) and speaking (Q-58) assessments are given. According to the results, teachers generally use feedback to highlight students' strengths and weaknesses and offer suggestions for improvement.

The most popular feedback strategy for listening tests chosen by (Q-24) is “*Highlighting students' strengths and weaknesses*” (37.1%). This percentage suggests that teachers view listening tests as a tool to guide student learning, as well as a method to assess their listening skills. Furthermore, 30.1% of the teachers stated that they prefer “*By providing suggestions for improvement*” to offer suggestions for feedback and improvement. This demonstrates that student-centered assessment is embraced. Although it is a less popular approach, student-teacher conferences (25.9%) demonstrate that personalized feedback is a crucial practice area. However, “*Preparing development reports*” (7%) was the least popular approach, indicating that listening assessments do not typically provide comments based on official documents.

There was a comparable feedback technique for speech assessment (Q-58). Once more, “*Highlighting students' strengths and weaknesses*” (32.2%) was the approach that teachers favored the most; however, this percentage was just lower than that of the listening tests. Furthermore, 31.5% of the teachers reported using it to offer recommendations for better feedback, a rate that closely resembles that of listening tests. In speaking assessments, student-teacher conferences were used more frequently (26.9%), suggesting that individual feedback may be more important for improving speaking skills. In contrast to listening tests, the approach of “*preparing development reports*” (10.5%) had a higher rate, indicating that more thorough documentation may be required for evaluating speaking abilities.

The feedback techniques used in speaking and listening tests exhibited a significant degree of statistical consistency. Direct feedback appears to be a more successful approach to improving oral abilities, as seen by the increased preference for individual student interviews in speaking assessments. However, the growing use of development reports in speaking evaluations indicates an attempt to track the long-term evolution of students' oral ability in greater depth.

These results show that teachers value giving students feedback on speaking and listening tests. Student feedback is most frequently used to pinpoint strengths and shortcomings and offers directions for improvement. However, teachers are approaching the process of evaluating speaking skills from a long-term perspective, as seen by the growing emphasis on individual interviews in speech evaluations and the regular writing of improvement reports.

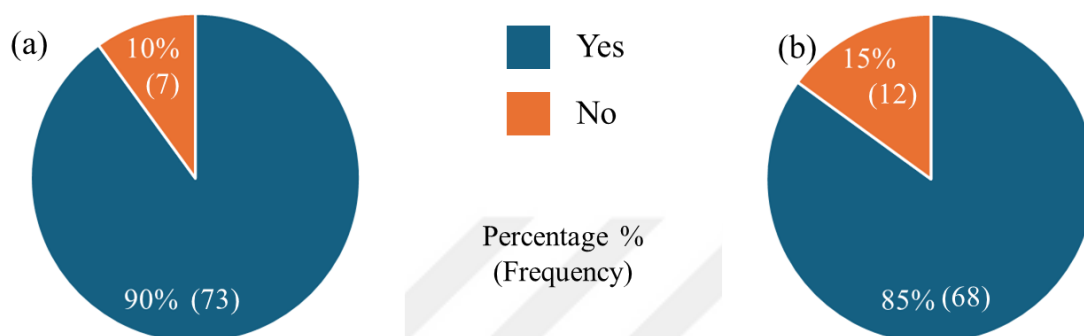
These findings imply that teachers might find it advantageous to employ resources such as digital feedback platforms, individual assessment sessions, and structured feedback frameworks more frequently to improve the efficiency of their assessment procedures. It may be advisable to

use standardized assessment criteria and feedback systems that can track students' progress over time, particularly when evaluating speaking abilities.

Figure 11: Statistics and Frequency Values for Q-25 & Q-59

Item-25: Do you use student feedback in the **listening** test assessment process?

Item-59: Do you use student feedback in the **speaking** test assessment process?



The statistical findings presented in Figure 11 reveal English teachers' perceptions of the extent to which they integrated student feedback into the assessment process for listening (Q-25) and speaking (Q-59) evaluations. The results indicate that student feedback is widely used in both listening and speaking assessments, but this integration is somewhat more limited in speaking assessments.

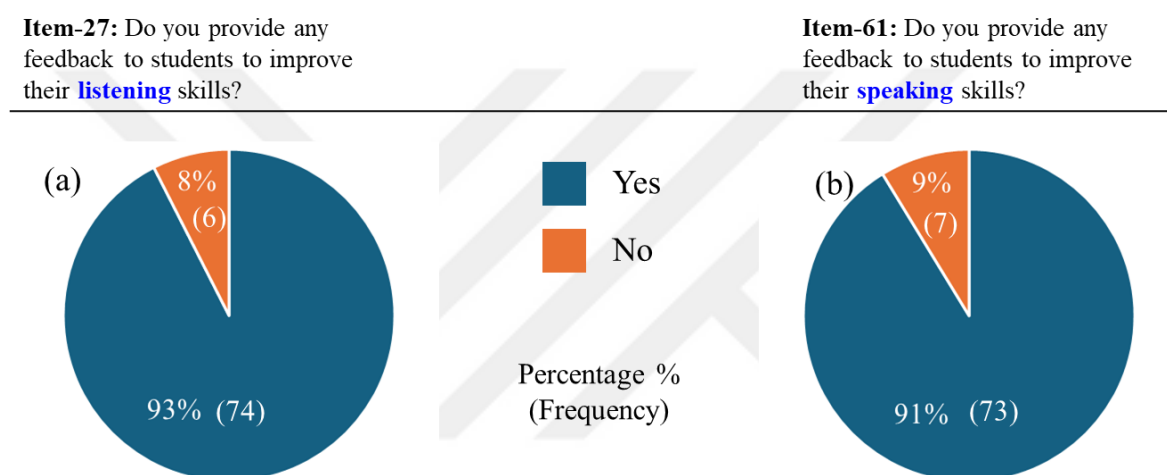
In listening assessments (Q-25), 90% of teachers indicated that they included student feedback in their evaluation processes. Procedures are considered the input they receive from the students' exam experiences. Given that only 10% of teachers did not incorporate student responses into the evaluation process, it appears that taking student comments into account when administering listening tests is a standard procedure.

Similar trends were seen in speaking assessments (Q-59), although a slightly smaller percentage of student comments (85%) was incorporated into the assessment process. This implies that incorporating student input into speaking assessments may be more challenging for teachers. Including feedback from students in the scoring process can be challenging, especially when oral assessments are subjective. Additionally, 15% of teachers for speaking assessments did not incorporate student feedback into the evaluation process, a slightly larger percentage than the 10% for listening examinations.

According to statistics, although it is standard practice to incorporate student feedback into assessment procedures for both skills, speaking tests continue to have a somewhat lower rate of integration. The reason for this discrepancy is that listening tests are more structured, facilitating

the incorporation of student comments. However, because the evaluation criteria are subjective, it can be more challenging to incorporate student comments into the speaking assessment scoring procedures. Although this occurs at a somewhat lower percentage in speaking tests, these results generally show that professors are willing to use student comments in evaluation procedures. To better include student input in assessment procedures, structured feedback systems can be created. This procedure can be made more efficient by providing students with one-on-one feedback sessions following tests, using digital feedback platforms, or organizing the feedback process through peer reviews.

Figure 12: Statistics and Frequency Values for Q-27 & Q-61



The statistical findings presented in Figure 12 reveal the tendency of English teachers to provide feedback to students in the process of developing listening (Q-27) and speaking (Q-61) skills. The results indicate that the vast majority of teachers view feedback as an important component of the language learning process and widely adopt it.

According to the listening skills evaluations (Q-27), 93% of the teachers said that they provided their students with feedback. This rate demonstrates that the evaluation process is employed as both a learning-oriented and a result-oriented instrument, highlighting the significance of feedback in the development of listening skills. However, only 8% of the respondents said they do not provide students' comments, indicating that giving feedback is standard practice in listening instruction.

Assessments of speaking ability showed a similar pattern (Q-61). According to 91% of the teachers, they provided their students with feedback on their speaking abilities. However, compared with speaking skills exams, a significantly higher percentage of teachers (9%) did not offer feedback on speaking skills assessments. This minor variation implies that speaking evaluations require a more intricate feedback procedure from teachers. It can be particularly challenging for

teachers to provide feedback in speaking evaluations due to the lengthy nature of the oral performance feedback process and the inclusion of subjective evaluation components.

According to statistics, teachers are more likely to provide students with comments on their speaking and listening abilities. However, the somewhat lower feedback rate in speaking tests suggests that teachers find this area's feedback procedures more difficult. Although the speaking skills feedback procedure is more challenging, these findings generally demonstrate that teachers are committed to providing feedback to their students. To enhance the effectiveness of feedback in speaking assessments, teachers are encouraged to utilize organized assessment rubrics, offer individual feedback sessions to students, and implement digital feedback tools.

Based on these results, teachers need to utilize diverse methods to effectively manage the enhancement of speaking and listening skills, as feedback plays a crucial role in this process. Improving the pertinence of feedback, especially regarding speaking skills, may help teachers to support children's language growth more effectively.

Table 14: Statistics and Frequency Values for Q-28 & Q-62

Questions	Q28. What feedback do you provide to students to improve their listening skills?		Q62. What feedback do you provide to students to improve their speaking skills?	
Options	Frequency	%	Frequency	%
Correcting grammatical and pronunciation errors	15	20,3	18	24,3
Evaluating content and meaning	27	36,5	15	20
Focus on speaking fluency and expression skills	18	24,3	22	30,3
Encouraging students to build on their strengths	11	14,9	19	25,4
Other	3	4,1	0	0
Total	74	100	73	100
Statistics	Mean	2,46	Mean	2,57
	Median	2	Median	3
	Mode	2	Mode	3

The statistical findings presented in Table 14 reveal the tendencies of English teachers regarding the content of feedback they provided to students during the evaluation process of listening (Q-28) and speaking (Q-62) skills. According to the results, feedback on listening skills is focused on meaning and content. However, feedback on speaking skills is concentrated on fluency and self-expression abilities.

In listening assessments (Q-28), “*Evaluation of content and meaning*” (36.5%) is the most frequently preferred feedback type by teachers. This means that teachers, to improve listening skills, focus on the extraction of meaning and comprehension. In addition, 24.3% of them stated that they gave listening feedback by “*focusing on speaking fluency and expression skills*”. This leads to the idea that listening assessments can be evaluated by integrating listening skills with speaking. On the other hand, “*Correction of grammar and pronunciation errors*” (20.3%) was a less preferred type of feedback, which may mean that listening skill assessments are more focused on general comprehension and communication.

In speaking assessments (Q-62), the focal point of the feedback changed significantly. The most frequently preferred type of feedback by teachers was “*Focusing on speaking fluency and expression skills*” (30.3%), indicating that fluent and natural expressions were prioritized in speaking assessments. Furthermore, the greater percentage of “*Encouraging students to develop their strengths*” (25.4%) than in listening tests (14.9%) suggests that speaking skill exams use a strategy to boost confidence and motivation. However, “*Correcting grammatical and pronunciation mistakes*” (24.3%) also surfaced as a significant feedback factor in speaking skill assessments, indicating that pronunciation and grammatical accuracy are crucial elements in speaking skill evaluation.

Statistically speaking, it was observed that there is a greater variety of feedback in speaking assessments. While teachers adopt a meaning- and content-oriented approach in listening assessments, they follow a more comprehensive feedback process in speaking assessments to increase fluency, develop students' strengths, and correct grammatical errors.

In general, these findings reveal that teachers focus on content and meaning in listening skill assessments, whereas they focus on developing fluency, expression skills, and student motivation in speaking assessments. These differences can be explained by the fact that listening skills are receptive skills that require assessments to measure students' ability to understand and infer correctly. On the other hand, since speaking is a productive skill, it can be said that elements such as fluency, grammatical accuracy, and self-confidence gain more importance in the assessment process.

These results indicate that it may be beneficial for teachers to use structured feedback systems to make speaking assessments more effective. Besides, increasing feedback sessions focusing on fluency and expression skills may also be more beneficial.

Table 15: Statistics and Frequency Values for Q-29 & Q-63

Q29. Which criteria do you use when assessing students' <i>listening</i> skills?			Q64. Which criteria do you use when assessing students' <i>speaking</i> skills?		
Options	Frequency	%	Options	Frequency	%
Level of comprehension	58	16,8	Speech fluency	74	19
Understanding the main idea/topic	47	13,6	Grammatical accuracy	46	11,8
Understanding the details	25	7,2	Vocabulary richness	61	15,7
Vocabulary richness	53	15,4	Content (ideas, details)	46	11,8
Define key concepts/words	47	13,6	Pronunciation	62	15,9
Understanding the pronunciation of words	48	13,9	Interaction	32	8,2
Understanding intonation and emphasis	25	7,2	Body language (gestures and facial expressions)	39	10
Making inferences after listening	42	12,2	Voice tone (intonation)	28	7,2
Other	0	0	Other	1	0,3
Total	345	100	Total	389	100
Dichotomy group tabulated at value 1.					

Table 15 depicts the statistical data on the tendencies of EFL teachers regarding the criteria they prefer when assessing listening (Q-29) and speaking (Q-63) skills. Considering results, listening assessments focus on comprehension and vocabulary, while speaking assessments focus on fluency, pronunciation, and oral expression skills.

The most frequently used criterion in listening assessments (Q-29) was “Level of *comprehension*” (16,8%). This is because teachers give greater importance to measuring the degree to which students understand from listening assessments. Furthermore, “*Vocabulary richness*” (15.4%) and “*Understanding word pronunciation*” (13.9%) were among the often-used assessment criteria. This also indicates that listening skills measure vocabulary knowledge, pronunciation, and general meaning comprehension. “*Making inferences after listening*” (12.2%) was another criterion taken into consideration, which is the student's listening capacity. These indicate that listening assessments evaluate critical thinking skills along with comprehension.

In speaking assessments (Q-63), the distribution of the criteria differs. With “*Speaking fluency*” (19%), which indicates that students' linguistic fluency is seen as the most important feature in speaking assessments, teachers most often cited it as their criterion. Furthermore, highly valued by instructors were “*Pronunciation*” (15.9%) and “*Vocabulary richness*” (15.7%), thereby demonstrating that speaking skills are evaluated not only by fluency but also by word usage and pronunciation correctness. The grading criteria also included nonverbal elements of oral communication, such as “*body language (gestures and facial expressions)*” (10%) and “*tone of voice (intonation)*” (7.2%). However, these criteria were less important.

From a statistical point of view, vocabulary knowledge is seen as an important factor in both listening and speaking skill assessments; however, listening assessments focus on understanding word pronunciation, whereas speaking assessments focus more on the variety of words used. Similarly, although pronunciation is an important assessment criterion for both skills, teachers treat pronunciation as a direct production-based criterion in speaking assessments.

Generally, these findings reveal that teachers prioritize meaning level, vocabulary, and critical listening skills in listening assessments. In contrast, fluency, pronunciation, and vocabulary use are seen as the most critical elements in speaking assessments. This shows that listening skills are primarily associated with comprehension and information processing processes, whereas speaking skills are directly related to production and expression skills. Findings show that in speaking examinations, not only grammar accuracy but also techniques should be created to improve fluency and communicative competency. Providing students with more efficient feedback systems can help improve their communication abilities, particularly in areas such as pronunciation, fluency, and vocabulary richness.

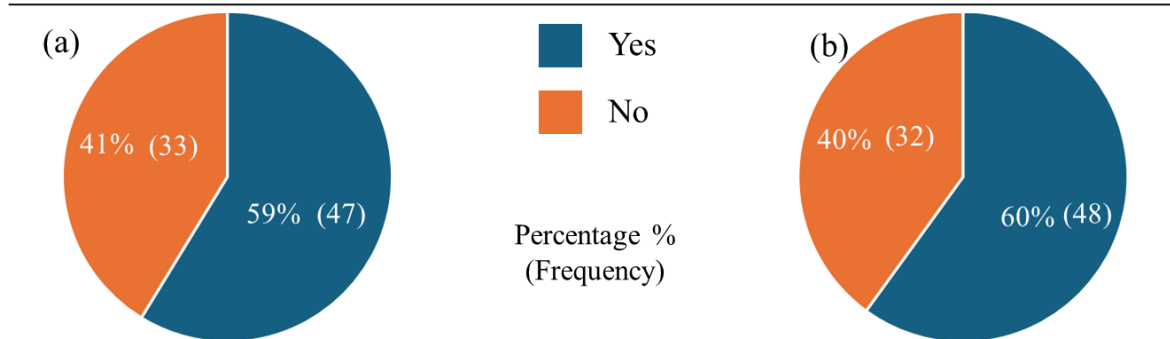
English teachers' opinions on whether they received sufficient instruction in the assessment of listening (Q-30) and speaking (Q-64) skills during their undergraduate study are statistically revealed in Figure 13. According to the findings, most teachers felt that they had sufficient instruction in assessing these abilities throughout their undergraduate studies; nonetheless, a sizable fraction also felt lacking in this area.

With reference to listening assessment training (Q-30), 59% of the teachers said that they received sufficient instruction in this field. According to this statistic, most undergraduate students believe they are proficient in conducting listening tests. Nonetheless, 41% of the participants reported not having received sufficient instruction in listening evaluation, suggesting that training in this field is not sufficiently thorough for all teachers.

Figure 13: Statistics and Frequency Values for Q-30 & Q-64

Item-30: Did you receive adequate training on assessing **listening** skills in your undergraduate education?

Item-64: Did you receive adequate training on assessing **speaking** skills in your undergraduate education?

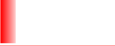


Examining speaking assessment training (Q-64), one can find a comparable pattern. Although 40% of the teachers said they lacked sufficient training in speaking assessment, 60% said they had sufficient training in this regard. This percentage mostly corresponds with training in listening assessment. Therefore, teacher preparation schools offer a comparable degree of education in both listening and speaking. Statistically, the teachers' opinions on speaking and listening assessment training were quite similar. About 40% of teachers stated they did not have enough knowledge in both areas. This shows that undergraduate programs should teach assessment skills thoroughly. This rate suggests that assessing language should be a priority in teacher training courses.

These results show that teachers received limited training in assessing listening and speaking skills during undergraduate education, but this is not sufficient for all of them. Almost 40% of teachers state they need more training in assessing skills. This means better and more practical teaching methods are required in both areas. All these suggest more topics for the undergraduate curriculum, such as assessment techniques and structured rubrics, so that teacher candidates get better assessment training through their undergraduate education. Particularly in fields where subjective elements, including speaking evaluation, are more noticeable, extra training courses and seminars might be advised to assist teachers in better and more consistent management of assessment procedures.

The statistical findings presented in Figure 14 reveal EFL teachers' perceptions of the quality of the training they received during their undergraduate education regarding the assessment of listening (Q-31) and speaking (Q-65) skills. The results showed that listening assessment training was evaluated more favorably than speaking assessment training.

Figure 14: Statistics and Frequency Values for Q-31 & Q-65

Item-31: How do you evaluate this training you received in your undergraduate education for the assessment of listening skills?			Item-65: How do you evaluate this training you received in your undergraduate education for the assessment of speaking skills?		
Options	Percentage	Frequency	Percentage	Frequency	
<i>Very good</i>	 39%	31	 15%	12	
<i>Good</i>	 15%	12	 30%	24	
<i>Fair</i>	 31%	25	 10%	8	
<i>Insufficient</i>	 13%	10	 5%	4	
<i>Very insufficient</i>	 3%	2	 40%	32	
Statistics	Mean:	2.2500	Mean:	3.2500	
	Median:	2	Median:	3	
	Mode:	1	Mode:	5	

Regarding listening assessment training (Q-31), 15% of the teachers assessed it as “good,” and 39% of the teachers assessed it as “very good.” These results reveal that a sizable fraction of teachers believe that their undergraduate degree provides adequate listening assessment instruction. However, 31% of respondents said that the training was “moderate,” implying that it was not thorough enough for every teacher. Furthermore, it was revealed that some teachers sensed shortcomings in this area, as 13% assessed the training as “inadequate” and 3% as “very inadequate.”

When evaluated in terms of speech evaluation training (Q-65), the teachers' evaluations of their training in this area were more negative. Only 15% of respondents said the instruction was “very good”, while 30% said it was “good”. Still, 40% said the instruction was “moderate”, 10% said it was “insufficient”, and 5% said it was “very insufficient”. These figures show that the degree of satisfaction with speaking assessment training was less than that with listening assessment training. More favorable opinions of hearing assessment training ($M = 2.25$) were indicated statistically by its average satisfaction score being lower than that of the speaking assessment training ($M = 3.25$). Furthermore, the mode value for speech assessment training was five (very poor), indicating that many teachers thought this training was insufficient. Generally speaking, these results reveal that although instructors felt their undergraduate training in listening evaluation was better, they felt their training in speaking assessment was insufficient. Many teachers said that their speech assessment training was either insufficient or limited in scope.

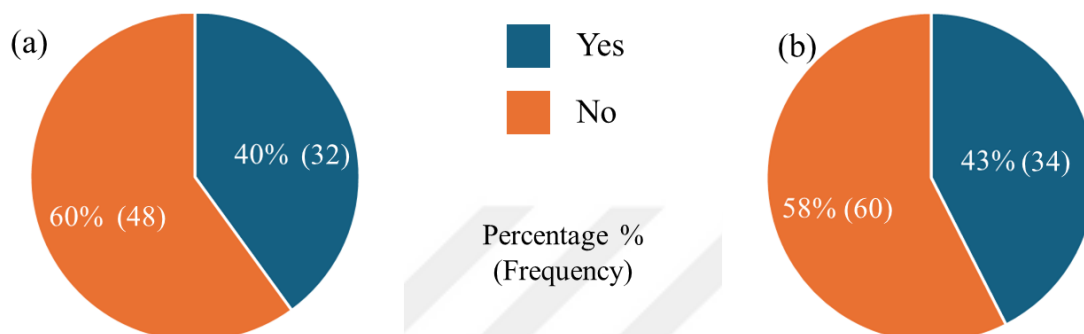
The results indicate that speech assessment training is more effective. Specifically, to enhance the reliability of speech assessment processes, it is recommended that structured assessment criteria, objective scoring methods, and practical applications aimed at improving teachers' assessment skills be integrated into undergraduate education. Increasing the number of such

training programs contributes to reducing the challenges teachers face when conducting speech assessments and to fostering more objective and consistent evaluation practices.

Figure 15: Statistics and Frequency Values for Q-32 & Q-66

Item-32: Have you participated in any professional activities (conferences, seminars, webinars, etc.) to assess **listening** skills after undergraduate studies?

Item-66: Have you participated in any professional activities (conferences, seminars, webinars, etc.) to assess **speaking** skills after undergraduate studies?



The statistical findings presented in Figure 15 reveal the tendencies of English teachers regarding their participation in professional development activities aimed at the assessment of listening (Q-32) and speaking (Q-66) skills after their undergraduate education. The results showed that the vast majority of teachers did not receive any professional training in these areas.

Regarding listening assessment training (Q-32), just 40% of the teachers said that they engaged in professional development activities in this field following their undergraduate degree. Conferences, seminars, and online training courses are among these pursuits. However, 60% of instructors said they had not received any further training in listening assessment, implying either limited possibilities for professional growth in this field or teachers not making sufficient use of available opportunities.

Analyzed in terms of speaking assessment training (Q-66), a similar tendency was found; the rate of obtaining training in this area was somewhat greater. Although 57% of the teachers reported not having engaged in professional development activities in the field of speaking assessment, 43% stated that they had. This slight difference implies that teachers might have access to greater professional development opportunities in this field or obtain more training to enhance their speaking assessment skills.

Statistically, the participation rates in both listening and speaking assessment training were relatively low. However, the participation rate in the speaking assessment training (43%) was slightly higher than that in the listening assessment training (40%). This is because speaking

assessments are more subjective, and teachers require additional training to improve their assessment skills.

In general, these findings indicate that listening and speaking assessments are not integrated into professional education during undergraduate studies. This may lead to increased difficulties for teachers' assessment processes and prevent them from assessing students' development. That is why, teachers should access more professional development trainings such as applied trainings and certificate programs.

Table 16: Statistics and Frequency Values for Q-33 & Q-67

Questions	Q33. Would you like to participate in a training on assessing listening skills? How useful do you think such an event would be?		Q67. Would you like to participate in a training on assessing speaking skills? How useful do you think such an event would be?	
Options	Frequency	%	Frequency	%
Yes, I would like to participate, and I think it would be very useful	27	33,8	26	32,5
Yes, I would like to participate, and I think it would be useful.	32	40	32	40
I am undecided.	14	17,5	13	16,3
No, I would not like to participate, but I think it would be useful.	3	3,8	6	7,5
No, I do not want to participate, and I do not think it will be useful	4	5	3	3,8
Total	80	100	80	100
Statistics	Mean	2,0625	Mean	2,1
	Median	2	Median	2
	Mode	2	Mode	2

The statistical findings presented in Table 16 detail the willingness of EFL teachers to participate in training programs covering the assessment of listening (Q-33) and speaking (Q-67) skills, and their opinions on whether such training is proper or not. The results showed that the majority of the teachers wanted to receive professional training in both listening and speaking assessments and believed that such training would be helpful.

In terms of listening assessment training (Q-33), 73.8% of teachers stated that they would like to receive training in this field. This reveals that teachers show great interest in professional development opportunities to improve their listening assessment skills. While 33,8% of this group thought that this form of training would be “*very useful*,” 40% responded that they would find it “*useful*.” However, 17.5% of the respondents were unsure about the substance of the training courses or their efficacy, implying that some teachers were not sure about either. Conversely, 5% said they were totally uninterested in listening evaluation training, while 3.8% stated that they would find the training helpful but would not participate in it.

When analyzed in terms of speaking assessment training (Q-67), a similar trend was observed. 72,5% of the teachers stated that they would like to receive training in this field, which largely parallels the rate stated for listening assessment training. While 32,5% of this group thought that speaking assessment training would be ‘*very useful*,’ 40% stated that they would find it ‘*useful*.’ The rate of those who were undecided was 16.3%, which is close to the rate of indecision regarding listening assessment training. Additionally, 7.5% of the respondents agreed that this training would be useful but stated that they would not participate. In comparison, 3.8% of the respondents were completely uninterested in this kind of training.

From a statistical perspective, teachers' willingness to participate in both listening and speaking assessment training was relatively high. However, interest in listening assessment training (73.8%) was slightly higher than that in speaking assessment training (72.5%). The fact that this difference is relatively small suggests that teachers require professional development for both skills, indicating a similar level of importance.

Generally, these findings reveal that teachers are willing to improve their assessment of listening and speaking skills, but some teachers are uncertain about the effectiveness of such training. Tailoring such professional development opportunities to the needs of teachers and offering them effectively can contribute to further development of teachers' assessment skills. Findings support the need for teachers to receive more training in listening and speaking assessments. Practical training can help teachers learn how to make their assessment processes more effective in real classroom settings. Moreover, the design of structured training programs to address teachers' doubts may encourage undecided teachers to participate.

The statistical findings presented in Table 17 reveal EFL teachers' preferences for the types of professional activities that they find most productive in enhancing their professional development in the assessment of listening (Q-34) and speaking (Q-68). According to Table 17, “Seminars” (22,7%) are the most preferred option in both assessment skills. However, speaking assessment training is thought to be supported by more interactive activities.

Table 17: Statistics and Frequency Values for Q-34 & Q-68

Questions	Q34. Which of the following professional activities for assessing listening skills would be more productive for you at the moment?		Q68. Which of the following professional activities for assessing speaking skills would be more productive for you at the moment?	
	Frequency	%	Frequency	%
Conferences	29	13,2	33	13,3
Seminars	50	22,7	51	20,6
Webinars	35	15,9	31	12,5
Panel Discussions	14	6,4	25	10,1
Presentations	32	14,5	38	15,3
Debates	8	3,6	20	8,1
Education Competitions	15	6,8	18	7,3
Online Courses	37	16,8	32	12,9
Total	220	100	248	100
Dichotomy group tabulated at value 1.				

In listening assessment training (Q-34), “*Seminars*” (22.7%) are the most preferred professional activity by English teachers. This shows that teachers believe that structured discussions with experienced leaders are the most effective method for developing listening assessment skills. In addition, “*Online courses*” (16.8%) and “*Webinars*” (15.9%) were the other professional development formats commonly preferred by teachers. These suggest that digital and self-paced learning activities are particularly valuable for developing listening assessment skills.

Considering speaking assessment training (Q-68), it is seen that teachers prefer “*Seminars*” (20,6%) the most. It is noteworthy that, in this skill area, activities involving more interaction are preferred. “*Presentations*” (15,3%) and “*Panel discussions*” (10,1%) were the preferred activities by teachers in speaking assessments. As can be understood from this, teachers think that speaking assessment training should not be limited to individual knowledge transfer but should also be based on discussion and experience sharing. Statistically, teachers find structured learning formats (such as seminars and online courses) to be the most useful for both listening and speaking assessment training.

In general, teachers need different learning methods in their professional development process to enhance their language skills. As for listening assessment skills, structured and self-learning-oriented activities are used. On the other hand, the need shifts to more interactive and practical activities for speaking assessment skills. Therefore, listening and speaking assessment training requires different methods from each other to support the professional development of teachers effectively. Primarily, teachers should be provided with training programs including panel discussions, interactive workshops, and practical applications. Additionally, supporting teachers in

attending more online courses and seminars on listening assessment training may help them enhance their competencies in this field.

Table 18. Thematic Coding Table

Theme	Code	Participants
Theme 1: Contributions to Learners	Increased confidence	A, B, C, D, E, F
	Improved pronunciation	A, B, C, D
	Enhanced listening comprehension	A, B, C, F
	Better self-expression	A, B, C, D
	Transition from passive to active role	D
	Familiarity with exam types	F, D
Theme 2: Challenges and Disadvantages	Exam-related anxiety	A, B, C, D, E, G
	Limited vocabulary	A, B, E, D
	Technical limitations in classrooms	C, D, E
	Time constraints	A, B, D, E, G
	Increased workload for teachers	E, D, G
	Students refusing participation	C, D, E
Theme 3: Teacher Observations	Holistic student assessment	A, B, C, D
	Integrated view of language skills	C, D
	Spontaneity and use of body language	C
	Value of drama-based activities	D, G
Theme 4: Educational Impact	Encouragement through task-based work	D, C
	Increased learner awareness	A, D
	Pedagogical shift in assignments	C, D
Theme 5: Views on Assessment Methods	Process-oriented evaluation	D, G
	Speaking as a more valid measure	C, D, F
	Technical improvements needed	C, D
	Opposition to scripted tasks	D, G
	Teacher training needs	D
Theme 6: General Attitudes & Suggestions	Supportive views	A, B, C, D, F
	Mixed but constructive views	E
	Critical/opposing views	G
	Assessment should be integrated.	D, E, G
	Need for piloting and planning.	D

3.2. Qualitative Data

Interviews with the participants were evaluated, and thematic analysis was carried out. The themes were developed with expert opinions. In this section, the findings based on teacher and student opinions are presented under six main themes. These themes were categorized as the contributions of exams for students, difficulties and disadvantages encountered, teachers'

observations, educational contributions, opinions on assessment and evaluation methods, general attitudes, and suggestions (Table 18).

Table 19. Codes and Quotes for Theme 1: Contributions to Students

Code	Direct Quote
Increased Confidence	Participant A: When students know they will be assessed through speaking and listening, they take risks. I can see their self-confidence grow every week, especially when they realize they can express their ideas in English without first thinking in Turkish.
Improved pronunciation	Participant B: Repeated speaking tasks make a difference. I have noticed that my students start pronouncing complicated sounds more clearly after a few weeks, especially /th/ and /w/, which they previously avoided.
Enhanced listening comprehension	Participant F: Listening to exams pushes them to really focus. Before, they just skimmed dialogues in the book, but now they listen for meaning and details. This has improved their overall comprehension and note-taking skills.
Better self-expression	Participant C: I can see them becoming storytellers, not just test-takers. Even weaker students can describe their weekend or express an opinion. This is a significant change compared to before.
Transition from passive to active role	Participant D: They stop being passive listeners. Speaking tasks force them to take an active role in the lesson, and they begin to see English as a fundamental communication tool.
Familiarity with exam types	Participant D: Because they see these tasks in class, they are no longer shocked in the real exam. They know what to expect and how to perform.

Theme 1: Contributions to Students:

The participants emphasized that listening and speaking exams had multifaceted positive effects on students. The data indicate that listening and speaking exams contribute to students' learning processes in many ways. These contributions can be categorized under several sub-codes (Table 19). One of the most frequently mentioned benefits was increased self-confidence. Another highlighted contribution is the improvement of pronunciation. Similarly, the development of listening comprehension skills was also seen as an important outcome. Another contribution from the exams was better self-expression. This contribution is in line with the other one, which is a transition from passive to active role. Finally, familiarity with the types of exams was also highlighted as a contribution. The following statements support the obtained sub-codes under Theme-1: Contributions to Students.

When students know they will be assessed through speaking and listening, they take risks. I can see their self-confidence grow every week, especially when they realize they can express their ideas in English without first thinking in Turkish. (Participant A)

Repeated speaking tasks make a difference. I have noticed that my students start pronouncing complicated sounds more clearly after a few weeks, especially /th/ and /w/, which they previously avoided. (Participant B)

Listening to exams pushes them to really focus. Before, they just skimmed dialogues in the book, but now they listen for meaning and details. This has improved their overall comprehension and note-taking skills. (Participant F)

I can see them becoming storytellers, not just test-takers. Even weaker students can describe their weekend or express an opinion. This is a significant change compared to before. (Participant C)

They stop being passive listeners. Speaking tasks force them to take an active role in the lesson, and they begin to see English as a fundamental communication tool. (Participant D)

Because they see these tasks in class, they are no longer shocked in the real exam. They know what to expect and how to perform. (Participant D)

Generally, as shown in Table 19, listening and speaking exams are not only an assessment tool but also an element of the learning process that contributes to language development, self-confidence, and active communicative participation.

Theme 2: Challenges and Disadvantages

This theme covered the structural and pedagogical barriers encountered in the implementation of exams. As seen in Table 20, one of the most frequently stated problems is test anxiety. Another difficulty is related to limited vocabulary knowledge. Technical problems encountered in practice also drew attention. In addition, time constraints also emerged as an important obstacle. Finally, the increasing workload of teachers is a notable disadvantage. Statements of participants supporting these sub-codes are given below:

Some students freeze. They know the vocabulary but forget everything when speaking in front of the teacher. This anxiety affects their real performance. (Participant E)

Even motivated students cannot always find the right word under pressure. They end up using Turkish or remaining silent, which unfairly lowers their scores. (Participant A)

The sound system in many classrooms is outdated. Some students cannot hear clearly from the back. It is not fair, but we cannot control it. (Participant C)

Testing everyone properly takes hours. With large classes, you cannot give each student enough time to speak and listen carefully. (Participant B)

We are preparing, grading, and recording three different exams for one course. It is exhausting, and it is not recognized as extra workload. (Participant E)

Participant D: Some simply refuse to speak. They say they are shy or not ready. It is a challenge to encourage them while keeping the flow of the exam.

While listening and speaking tests have significant benefits for students, the challenges summarized in Table 20 make it challenging to ensure their sustainability and fairness.

Table 20. Codes and Quotes for Theme 2: Challenges and Disadvantages Theme

Code	Direct Quote
Exam-related anxiety	Participant E: Some students freeze. They know the vocabulary but forget everything when speaking in front of the teacher. This anxiety affects their real performance.
Limited vocabulary	Participant A: Even motivated students cannot always find the right word under pressure. They end up using Turkish or remaining silent, which unfairly lowers their scores.
Technical limitations in classrooms	Participant C: The sound system in many classrooms is outdated. Some students cannot hear clearly from the back. It is not fair, but we cannot control it.
Time constraints	Participant B: To test everyone properly takes hours. With large classes, you cannot give each student enough time to speak and listen carefully.
Increased workload for teachers	Participant E: We are preparing, grading, and recording three different exams for one course. It is exhausting, and it is not recognized as extra workload.
Students refusing participation	Participant D: Some simply refuse to speak. They say they are shy or not ready. It is a challenge to encourage them while keeping the flow of the exam.

Table 21. Codes and Quotes for Theme 3: Teachers' Observations Theme

Code	Direct Quote
Holistic student assessment	Participant B: This way, we see their real abilities. Sometimes, a student who fails written grammar tests surprises me with their fluency in speaking tasks.
Integrated view of language skills	Participant C: You can observe listening, speaking, and even reading together. For example, when they listen to a prompt and then respond orally.
Spontaneity and use of body language	Participant D: Students use gestures, facial expressions, and tone. It shows understanding beyond words and helps shy students communicate.
Value of drama-based activities	Participant G: Drama brings English alive. Even my most introverted students take a role and perform. It is less threatening than formal speaking tests.

Theme 3: Teachers' Observations:

Under this theme, teachers stated that thanks to these exams, they had the opportunity to evaluate students not only in writing, but also in many ways. It was stated that improvisation, gesture-mimicking, and body language reflected the competence of students. Activities such as drama and role-playing increased the participation of low-performing or timid students in the process.

Table 21 shows that listening and speaking exams provide teachers with the opportunity to observe students in a multifaceted, natural, and holistic way. Statements of participants proving these sub-codes under Theme-3: Teachers' Observations are given as follows:

This way, we see their real abilities. Sometimes, a student who fails written grammar tests surprises me with their fluency in speaking tasks. (Participant B)

You can observe listening, speaking, and even reading together. For example, when they listen to a prompt and then respond orally. (Participant C)

Students use gestures, facial expressions, and tone. It shows understanding beyond words and helps shy students communicate. (Participant D)

Drama brings English alive. Even my most introverted students take a role and perform. It is less threatening than formal speaking tests. (Participant G)

Table 22. Codes and Quotes for Theme 4: Educational Contributions Theme

Code	Direct Quote
Encouragement through task-based work	Participant D: When I design task-based activities, students feel it is more like a game than an exam. They participate more actively and improve naturally.
Increased learner awareness	Participant A: They begin to realize language is a tool for communication, not just a subject. They notice their progress, which motivates them to study more.
Pedagogical shift in assignments	Participant C: I now assign listening to podcasts and speaking journals instead of only worksheets. This shift has changed classroom dynamics completely.

Theme 4: Educational Contributions:

The implementation of listening and speaking exams has led to changes in teaching methods. It was stated that students were motivated, and homework assignments were diversified in terms of content and method. That language is not only a written skill but also a multidimensional means of communication. One of the most important elements that encouraged more active student participation was task-based activities. Additionally, exams seem to have raised students' language awareness. In addition, there has been a pedagogical transformation in homework and classroom practices. The sub-codes and corresponding quotes from participants are detailed in Table 22. It shows that listening and speaking exams make an educational contribution to the teaching process not only in terms of assessment but also by renewing pedagogical approaches.

When I design task-based activities, students feel it is more like a game than an exam. They participate more actively and improve naturally. (Participant D)

They begin to realize language is a tool for communication, not just a subject. They notice their progress, which motivates them to study more. (Participant A)

I now assign listening to podcasts and speaking journals instead of only worksheets. This shift has changed classroom dynamics completely. (Participant C)

Theme 5: Assessment and Evaluation Opinions:

This theme discusses teachers' perspectives on how listening and speaking exams should be measured and assessed. As shown in Table 22, most participants argued that process-oriented assessment methods should be used instead of single-session exams. Additionally, the significance of speaking as a valid assessment tool was highlighted. However, technical improvements are needed to make implementation fairer and more effective. Some teachers also criticized the limitations of rote-based tasks. Finally, there was an emphasis on supporting teachers in the assessment process. Table 23 indicates that the assessment and evaluation of listening and speaking tests need to be reconsidered, and that teacher support, along with adopting process-oriented and more valid methods, is essential. We should evaluate students over time, not just in a single stressful moment. This approach would provide a more accurate picture of their abilities.

Written tests do not show who can use the language. Speaking reveals communicative competence much better. (Participant F)

Simple things like individual headphones or quieter rooms would make listening tests fairer for all students. (Participant C)

Scripted monologues feel artificial. Students memorize and forget. Spontaneous speaking gives us a more authentic measure. (Participant G)

We need workshops and clear rubrics. Assessing speaking fairly is much more complicated than marking a grammar test. (Participant A)

Theme 6: General Attitudes and Suggestions:

This theme includes teachers' general attitudes towards the implementation of listening and speaking tests and suggestions for improvement. Teacher opinions are divided into three groups: those who are positive, those who are critical but constructive, and those who are entirely opposed to implementation. It was suggested that teachers should be involved in the process, that the necessary infrastructure should be provided, and that pilot applications should be conducted to improve the implementation. Some teachers have emphasized that these skills should be incorporated into the lesson process independently of the exam.

Table 24 shows that teachers' views on implementation varied, but that the process could be improved with more inclusive planning, teacher participation, and appropriate infrastructure support. The supporting statements of on these details are given as follows.

This reform is necessary. Students cannot learn English only through writing. Speaking and listening are the missing pieces. (Participant B)

It is a good idea, but the timing and resources were not ready. We are building the plane while flying it. (Participant E)

We were not consulted. The system was forced on us without proper training or equipment. (Participant G)

Oral skills should be part of everyday lessons, not a separate, scary test at the end. (Participant D)

They should have piloted it with a small group first. Now we are dealing with all the problems at once. (Participant C)

Table 23. Codes and Quotes Theme 5: Views on Assessment Methods Theme

Code	Direct Quote
Process-oriented evaluation	Participant D: We should evaluate students over time, not just in one stressful moment. This would give a more accurate picture of their abilities.
Speaking as a more valid measure	Participant F: Written tests do not show who can use the language. Speaking reveals communicative competence much better.
Technical improvements needed	Participant C: Simple things like individual headphones or quieter rooms would make listening tests fairer for all students.
Opposition to scripted tasks	Participant G: Scripted monologues feel artificial. Students memorize and forget. Spontaneous speaking gives us a more authentic measure.
Teacher training needs	Participant A: We need workshops and clear rubrics. Assessing speaking fairly is much more complicated than marking a grammar test.

Table 24. Codes and Quotes for Theme 6: General Attitudes & Suggestions Theme

Code	Direct Quote
Supportive views	Participant B: This reform is necessary. Students cannot learn English only through writing. Speaking and listening are the missing pieces.
Mixed but constructive views	Participant E: It is a good idea, but the timing and resources were not ready. We are building the plane while flying it.
Critical/opposing views	Participant G: We were not consulted. The system was forced on us without proper training or equipment.
Assessment should be integrated.	Participant D: Oral skills should be part of everyday lessons, not a separate, scary test at the end.
Need for piloting and planning.	Participant C: They should have piloted it with a small group first. Now we are dealing with all the problems at once.

This study examined English language teachers' perceptions of the inclusion of listening and speaking components in the national English language exams. Analysis of the interview data revealed six main themes, covering both the perceived benefits and challenges associated with this reform. The findings showed that most teachers recognized the pedagogical value of integrating oral skills into formal assessment processes; however, some teachers expressed concerns about implementation, practicality, and student diversity.

Teachers emphasized that the listening and speaking components have a positive impact on student engagement, self-confidence, pronunciation, and the development of communicative competence. Most participants reported that students who previously showed low engagement in written tasks demonstrated increased interest in language when they participated in oral activities. However, problems such as technical difficulties (e.g., inadequate classroom sound systems), exam-related stress, time constraints, and teacher workload are commonly mentioned. Participants expressed different views on how oral skills should be assessed, with most teachers expressing a preference for process-oriented and context-sensitive approaches over rote-based one-off tests. In general, teachers' views reflect an awareness of the value of communicative language assessment; however, this understanding is balanced by classroom logistics and structural constraints.

The findings of the study are in line with the basic principles of the CLT approach, which advocates the integration of the four basic language skills of listening, speaking, reading, and writing into the teaching and assessment processes. The participants' emphasis on the benefits of listening and speaking assessment supports the theoretical assumption that language learning is most effective through real-time and meaningful communication (Savignon, 2002).

The theme of increased learner confidence and more active participation is also consistent with previous research, which suggests that oral tasks can enhance learner autonomy by reducing emotional barriers (Littlewood, 2007). The drama and role-playing activities frequently mentioned by participants suggest constructivist pedagogies, in which students are positioned as co-constructors of meaning.

However, this study also highlights systemic and infrastructural limitations that hinder effective implementation. Technical inadequacies and the lack of standardized rating scales overlap with problems reported in similar contexts and point to the need for both hardware investment and teacher training. Participants' emphasis on formative and in-class integrated assessment also revealed the tension between policy-driven summative practices and teachers' practical knowledge.

Particularly, the diversity of teacher attitudes, ranging from optimism to resistance, highlights the need for teacher participation in policy reforms. Participant D (Rabia) stated that the introduction of top-down practices without pilot studies or teacher consultation could reduce the impact of well-intentioned initiatives. This view is in line with the literature on teacher agency, which views teachers as co-creators of change, not just implementers.

In conclusion, including listening and speaking skills in national exams has clear pedagogical benefits. However, the findings highlight the importance of careful planning, teacher support, and adaptable implementation methods for success. Future research could explore how these assessments affect students' attitudes or examine the long-term impact of oral exams on their language proficiency and motivation.

CHAPTER FOUR

4. DISCUSSIONS

This chapter interprets the quantitative and qualitative findings to explain how English teachers in Trabzon operationalize the assessment of listening and speaking skills, and why their practices take the shapes observed. We first synthesize the major empirical patterns—semester-based administration, comprehension-oriented formats for listening, interaction-oriented formats for speaking, higher perceived reliability in listening, and greater validity uncertainty and scoring variability in speaking—and consider these patterns through the lens of assessment principles (construct alignment, reliability–validity trade-offs, and practicality). We then relate the results to classroom constraints (time, materials, logistics), learner-affective factors (anxiety, engagement, motivation), and teachers’ assessment literacy and professional preparation.

Building on this synthesis, the discussion evaluates pedagogical implications for task design, scoring, and feedback, and proposes feasible adjustments that preserve authenticity while improving consistency—particularly in speaking assessment. We also outline capacity-building priorities (rubric design, rater calibration, task banks, and feedback routines) suggested by teachers’ expressed needs and professional development preferences. Finally, we acknowledge contextual limitations of the study and point to future work that could triangulate teacher self-reports with rater-agreement data, artifact audits, and classroom observations to test whether the proposed interventions enhance reliability, validity, motivation, and learning outcomes

4.1. Quantitative Data

The findings reveal a coherent yet differentiated assessment framework for listening and speaking in this context. Across domains—administration, design, scoring, feedback, and teacher preparation—patterns converge on a central theme: teachers value oral-skill assessment and implement it regularly, yet they adjust their assessment practices according to the specific affordances and constraints of each skill. This results in the observed distinction between comprehension-oriented listening practices and interaction-oriented speaking practices.

Different types of support are also needed: listening needs structured feedback sessions and materials, while speaking needs individual counselling and group work, which fits with the social and practice-heavy nature of oral production. Elshawa et al. (2017) and Oktaviani et al. (2024) also highlight that structured feedback supports listening improvement, while speaking benefits from

social and group-based practice. Monthly and especially weekly assessments remain comparatively infrequent, consistent with teachers' reports of time and logistical constraints. Yustina et al. (2021) and Sahyoni and Zaim (2017) obtained similar findings that limited time and classroom challenges make frequent speaking assessments difficult. The higher feasibility ratings for listening versus speaking, together with the higher management difficulty mean for speaking, indicate that the transaction costs of speaking assessments (scheduling, classroom management, turn-taking, evidence capture) are materially greater than for listening. In their research, Fauziah and Anisah (2022) concluded that speaking assessments require more coordination and organization than listening ones. This helps explain why teachers concentrate speaking assessments on more episodic, high-value events (e.g., presentations, role-plays) rather than frequent routine checks. Other studies also report that teachers prefer fewer but more meaningful oral tasks for practical reasons (Yustina et al., 2021; Elshawa et al., 2017).

Teachers' choices show strong construct-format coherence. Listening is mostly operationalized with blank-filling and comprehension questions that target recognition, parsing, and inference—capabilities that are straightforward to sample and score. Studies done by Buck (1992), Pahl and Schule (2022), and Peltekov (2022) report that formats measuring key listening processes like recognition and inference, helping to ensure construct validity. Speaking is predominantly assessed through presentations, role-plays, discussions, and Q&A, allowing spontaneous language, audience awareness, and extended discourse. This is consistent with research showing that these formats encourage spontaneous speech and authentic interaction, essential for evaluating real communicative ability (Choi et al., 2023; Miyane, 2020; Nalani et al., 2024). The central tendencies (means, modes, medians) show that these are not fringe preferences but common practices.

The change in focus from measuring skills (listening) to motivating students (speaking) is similar to the different levels of risk for the two tests. Listening assessments are often standardized and emphasize precise task-construct alignment, which ensures clearer scoring and higher validity (Borger, 2019). Listening tasks can be highly standardized, so precision in task-construct alignment moves to the foreground. Borger (2019) also states that such controlled formats allow teachers to assess recognition and inference more transparently and efficiently. Speaking tasks, conversely, are social and performative; anxiety, self-confidence, and classroom dynamics exert strong influence on performance. Research performed by Zuber and Conzelmann (2019) and May (2009) confirms that these social and emotional factors greatly affect students' speaking performance, making speaking assessment more complex and variable. This is also where the greatest reported challenge—student involvement—emerges for speaking, while listening challenges cluster around time management and material fit. Saptiany and Prabowo (2024) show that maintaining engagement during speaking tasks is difficult because of these affective pressures, whereas listening problems usually relate to timing and material suitability. The mixed “teacher-

friendliness” judgments further substantiate that speaking demands more classroom orchestration and post-hoc appraisal effort.

Perceptions of higher reliability in listening and greater validity uncertainty in speaking are consistent with the objectivity–subjectivity gradient. Listening assessments tend to produce more objective and reliable results because they rely on clear right/wrong responses and standardized formats (Joe et al., 2015; Zedelius et al., 2018). Listening items yield clearer right/wrong evidence, and teachers expect similar/same scores from peers at high rates. Teachers’ scoring consistency in listening has also been reported in other studies, which found strong agreement among raters (Huang et al., 2020). Speaking entails multidimensional constructs—fluency, pronunciation, lexis, coherence, and nonverbal features—making raters’ weighting schemes more variable. These multiple dimensions increase subjectivity, as raters may value different aspects differently depending on context and preference (Saeed et al., 2019; Huang et al., 2020; Cox & Davies, 2012). The reduced “same score” expectation and higher “lower score” anticipation in speaking point to inter-rater variance and threshold differences. These patterns justify the teachers’ call—implicit in improvement preferences—for clearer rubrics, rater calibration, and standardized scoring procedures in speaking.

High rates of using test results and considering student feedback across both skills suggest a well-established formative orientation. Elghotmy (2023), and Cobbold and Wright (2021) indicate that teachers’ use of feedback in both skills reflects a strong formative approach focused on improving learning rather than merely measuring it. Yet the content of feedback diverges in principled ways: listening feedback targets meaning and comprehension, while speaking feedback targets fluency/expression, strength-building, and error correction. This is supported by Elghotmy (2023) and Cobbold and Wright (2021), who note that listening feedback often helps learners interpret meaning and refine comprehension, while speaking feedback supports fluency and communicative competence through focused correction and encouragement. Support modalities also differ: listening depends on structured feedback sessions and materials, while speaking favours individual counselling and group work, which fits with the social and practice-heavy nature of oral production. These patterns indicate that teachers not only assess different constructs but also close the loop using skill-appropriate scaffolds. This is similar to Dewi et al. (2020), who highlight that teachers adapt feedback strategies to each skill’s specific needs, effectively closing the formative assessment loop.

Teachers perceived both assessments as motivating and instructionally useful, with speaking exerting a somewhat stronger motivational pull. This is consistent with the studies of Saidalvi and Samad (2019) and Lestari and Asari (2023), which reported that speaking tasks create stronger motivation because they provide authentic goals, peer interaction, and visible progress. This is pedagogically plausible: speaking assessments provide visible, authentic communicative goals, a

peer audience, and opportunities for self-efficacy, while listening to assessments, though essential, can feel more receptive for students. Saidalvi and Samad (2019) also found that peer feedback and interaction during speaking assessments enhance students' confidence and engagement, increasing their desire to improve. The mixed views and slightly higher negative ratings attached to speaking's instructional impact likely reflect performance pressure and the subjective nature of grading; this underscores the value of low-anxiety formats, transparent criteria, and feedback to transform assessment events into learning opportunities. Providing structured and clear feedback can reduce stress and make speaking assessments more constructive and motivating for learners (Abuzied & Nabag, 2023; Sheen, 2004; Siddiqui et al., 2022).

Teachers' criteria choices map cleanly to construct demands: listening emphasizes comprehension, vocabulary, pronunciation, and inference; speaking emphasizes fluency, pronunciation, vocabulary range, and, to a lesser degree, nonverbal cues. This confirms construct validity in use: what teachers say they value is what they design for and score. Vocabulary serves as a cross-skill pivot, functioning effectively in speaking (lexical availability and range) and efficiently in listening (recognition/pronunciation comprehension).

For both skills, the inclination for project-based and portfolio-based approaches represents a shift towards process-oriented, long-term learning data. Srikaew et al. (2015) describe these methods as a shift from isolated testing toward holistic, long-term assessment of learner progress. This is what teachers want to do: monitor students' progress over time, reduce performance noise from isolated incidents, and base assessments on more comprehensive data. Srikaew et al. (2015) confirm that these approaches help teachers track development and gain a fuller picture of student abilities. Peer reviews, recorded presentations, and self-evaluations that force you to reflect on yourself are examples of artefacts for speaking. Carefully selected assignments that span a variety of genres and skill levels could serve as objects for listening.

About 40% of the participants said they weren't well-prepared for college, even though most of them said they were. This was especially true for speaking tests, which got lower quality ratings and made people less happy (mean difference and modal "very poor" rating). Post-degree participation (PD) was modest (40–43%), yet willingness to engage is high ($\approx 73\%$ for listening; $\approx 72\%$ for speaking). Teachers prefer seminars and, for speaking, more interactive PD (presentations, panels), while listening development leans into online courses and webinars. This suggests a clear capacity-building pathway: offer sustained, practice-rich PD emphasizing rubric design, rater calibration, task prototyping, and feedback scripting, with flexible online components for listening and collaborative workshops for speaking.

4.2. Qualitative Data

The interview findings provide a clear picture of how teachers experience the use of listening and speaking exams in their classrooms. In general, they view these exams as useful and necessary, but at the same time, they face several difficulties while applying for them. Teachers' comments show that the exams help students improve in many ways, but some problems make the process harder than expected.

Many teachers said that the exams helped their students become more confident and active. Students began to participate in class more frequently and speak more fluently. Bepari and Imam (2025) also found that speaking assessments motivate students and increase their classroom participation and fluency. After completing speaking exercises multiple times, some of them started to listen more intently, while others saw an improvement in their pronunciation. Teachers observed that students gradually lost their fear of making mistakes and started speaking English more fluently. Muthmainnah et al. (2025) similarly report that supportive and interactive classroom environments help reduce students' fear of errors and promote freer communication. For the majority of teachers, this was one of the most significant shifts because students began to view English as a tool rather than as a topic they needed to learn.

Teachers stated additional challenges, although the overall impact was positive. Mpanza and Banja (2025) indicate that, despite positive outcomes, teachers still face notable difficulties during assessment implementation. One of the most common was exam anxiety. They said that some students know what to say but freeze when it is their turn to speak. Others cannot find the right word quickly and switch to Turkish. Such reactions are often linked to pressure under time limits, which can increase anxiety and cause language-switching behaviors (Tleuov, 2025). Time limits, big class sizes, and old classroom equipment were also mentioned often. Some teachers explained that listening exams are difficult to organize because of poor sound systems. These problems are consistent with previous research highlighting that large classes, poor acoustics, and outdated equipment make listening exams difficult to conduct effectively (Avedzi, 2025; Isakova & Kaldybaev, 2025). In addition, preparing and grading oral exams takes a lot of time and increases their workload. Mirsanjari (2025) similarly emphasizes that oral exam preparation and grading are time-consuming and add to teachers' overall workload. A few teachers added that some students refuse to take part because they are too nervous. These problems show that while teachers agree on the value of these exams, they need more support and better classroom conditions to manage them properly.

Teachers also explained that these exams allow them to see their students' real abilities. Unlike written exams, oral exams show how students use gestures, facial expressions, and tone of voice when they communicate. These non-verbal cues—such as gestures, facial expressions, and

voice tone—are essential indicators of real communication skills and enhance the validity of oral assessment (Ockey et al., 2014). Some teachers mentioned that drama and role-play activities help shy students join the lesson more easily. They said that such activities make students more comfortable and give a more complete view of their progress. Because of that, many teachers found oral exams more realistic and fairer in showing what students can actually do. Barakat (2023) also states that such interactive activities encourage participation, especially among shy learners, by reducing tension and improving expressiveness.

When talking about how the exams should be done, teachers mostly preferred continuous or process-based evaluation instead of one-time exams. They believed that following students' progress over time gives a clearer idea of their skills. Research by Mpanza and Banja (2025) also supports that tracking students' development over time helps reveal real performance trends and growth. Some of the teachers thought that speaking exams were more valid than written ones. Because they show how well students can use English in real communication. Many teachers admitted that grading fairly is a common challenge. They mentioned that using clear scoring rubrics, showing model performances, and attending short training sessions could make the process easier. This difficulty is frequently noted in research, as the subjective nature of oral exams makes consistency harder to maintain (Marzuki, 2024). Most teachers also favored speaking tasks that allow students to express themselves naturally rather than memorizing dialogues, as this approach better reflects real communication skills.

Overall, teachers had a positive attitude toward the inclusion of listening and speaking exams, yet their views differed about how the new system was carried out. While some saw it as an important move toward improving language learning, others felt it was introduced too quickly without enough preparation or guidance. Cheung and Wong (2012) also emphasize that successful reform depends on sufficient preparation and teacher involvement, which some educators felt was lacking. Several teachers added that oral skills should be practiced regularly in class instead of being tested only at the end of each term. They also recommended starting with pilot applications and involvement of teachers more actively in planning and decision-making. This aligns with Al-Shammari et al. (2019), showing that pilot programs and active teacher participation in decision-making improve implementation success and teacher acceptance.

In general, the interviews showed that teachers support the new assessment approach. Besides, they believe it helps students become more motivated, confident, and engaged in learning. At the same time, they face real classroom challenges, especially related to time, technical issues, and workload. Their views reflect both support for communicative assessment and the need for practical solutions. With more preparation, teacher training, and better classroom resources, these exams can become a more balanced and effective part of language assessment.

As an English teacher in a secondary school, I also faced similar patterns with my colleagues in speaking assessments. I also welcome this change, as do other Secondary School teachers, and I think that integrating assessment of speaking has many benefits for our students. They are much more motivated to speak, and they improve their skills better than before as they are aware that they must give more importance to speaking. However, we have many challenges, as implied by the results. First, time is really restricted to assessing the students properly if the assessment is restricted to 5- 6 minutes of presentations, dialogues, drama performances, etc. There are more than 40 students in many of the classrooms, and when the speaking exam is restricted to 5-6 minute performances, many students tend to memorize the responses and may show a level of proficiency that they may not show in a spontaneous speech. Therefore, I argue that evaluating speaking skills through a single performance, rather than through an assessment of the process, is neither effective nor valid. This approach yields unreliable results and fails to accurately represent the students' true communicative competence. Some students who study hard might not do well when their speaking is tested alone. Others who do not prepare well might get low scores because the exam is short. However, I have found that if we assess students in the process in every course with tasks, dialogues, drama performances, and pair work, they learn better and perform better in the relaxed atmosphere of the class. We can also have the opportunity to be aware of their individual talents. I agree with the teachers who are in favor of assessing authentic performance of the students, as I have seen so far that its advantages are higher than traditional and short, isolated exams. In this way, students can have the opportunity to adjust their speaking to different contexts, social issues, and situations with their performances, which increases their motivation and makes it more enjoyable.

In my view, integrating speaking assessment into everyday classroom activities is much more useful. For instance, when I used process-based assessment with the digital tool Class Dojo, I noticed that students were much more willing to speak because they could earn points that would help their eggs grow, which encouraged their motivation to participate in speaking. In addition, I experienced some logistical problems, like the other teachers in the research. Assessing approximately 40 students is challenging in both terms—time and process, leaving very limited time for feedback. In addition, group discussions and pair work may lead to unequal performance among students. Therefore, it may require the teachers to be careful while assessing because some students may talk less and some may talk more. In such situations, in-person assessment has some advantages for teachers. Finally, I can critically say that the participants of this research are quite right to say that we need standardized rubrics and training for assessing speaking because I think the lack of standardized rubrics and training results in variability in rubrics and assessment strategies that the English teachers use in Secondary Schools, and this also results in lack of validity and reliability of these exams.

CONCLUSION AND IMPLICATIONS

This chapter analyzes the quantitative and qualitative data collected for the study. It presents general findings regarding EFL teachers' perceptions of listening and speaking skill assessments, current practices, challenges they encounter, and the support needed. Based on these findings, various suggestions are offered for improving assessment and evaluation processes across pedagogical and structural levels.

Quantitative Data Conclusions

This study aimed to explore the perceptions, current practices, challenges, and support needs of EFL teachers in Turkey regarding the assessment of listening and speaking skills from a multidimensional perspective. The results from the quantitative data offer a detailed view of teachers' assessment processes for these skills and provide valuable insights into the communicative assessment approach, which is increasingly important in language teaching.

Actually, I observed similar things in my teaching and assessing context, I am aware of how crucial communicative approach and try to implement it with authentic tasks in my class, however, I feel complicated how effectively and well I can assess my students' performance appropriate for communicative approach in a great deal.

One of the main conclusions of the study is that listening and speaking assessments are conducted systematically, although teachers believe that assessing listening and speaking skills is pedagogically essential. The most common practice was found to be “termly” assessment. This suggests that teachers assess these skills only a limited number of times during the year. Time constraints, curriculum pressures, and practical challenges in assessing oral skills are reasons for the low percentage of “weekly” or “monthly” assessment practices. This reveals a gap between teachers' statements emphasizing oral skills in teaching and the realities faced in practice.

In my experience, this gap is a great problem as we are trying to adapt communicative teaching with exam requirements. We always have a curriculum to cope with and to follow and it restricts us to do what we hope. We often try to cope with time constraints and curriculum pressures while we are in favor of having enough time for each student to perform listening and speaking with authentic tasks.

Remarkable differences were also observed between how listening and speaking skills were assessed. It was observed that structured and controlled test formats, such as fill-in-the-blank and

multiple-choice questions, were preferred by teachers for evaluating listening skills. This suggests that objectivity is crucial when assessing listening skills and that the assessment process should be more practical. Conversely, in assessing speaking skills, teachers tend to use interactive and communication-based methods like presentations, role-playing, and discussions. This situation underscores the need for more productive and contextualized assessment methods for speaking skills, considering their inherent nature.

Because of time constraints, we try to find practical solutions; therefore, listening is much more practical for us, as it is more structured, and speaking is more authentic, but it requires much time, which policymakers should think carefully about. We, as teachers, are always in a dilemma of practicality and authenticity because of curriculum requirements, although we are in favor of communicative competence and performance

The findings also show that teachers move away from traditional exam techniques toward process-oriented assessment methods, including portfolio assessment, project-based assessment, and structured feedback processes, which are among the alternative approaches that teachers often prefer. Teachers found that portfolio systems, which enable monitoring of long-term development, especially speaking skills, are more effective. This trend highlights the importance of evaluating students' language development over time, rather than just their immediate performance.

Among the challenges in the assessment process, the most significant is the inability to maintain objectivity, especially when evaluating speaking skills. While listening skills can be measured using more structured test formats, making their assessment more objective and reliable, speaking skills tend to be more open to teachers' interpretation, leading to issues with consistency. Specifically, uncertainties about how to score aspects like fluency, pronunciation, vocabulary variety, and content richness can lower the reliability of the evaluation. In this context, teachers noted that different evaluators may score speaking tasks differently and emphasized that assessment and evaluation practices in this area need greater standardization.

The impact of student motivation on assessment processes is a notable finding of this study. Although speaking exams have greater potential to boost students' motivation, it was also observed that these exams present more implementation challenges for teachers. Therefore, developing motivational yet practical exam formats is an essential requirement.

I have also frequently seen that role plays, dramas, discussions and etc., supported student motivation and well-being, but I also observed that better results could be achieved with smaller groups

The study also showed that the assessment training teachers received at the undergraduate level was largely insufficient. Although they were provided with theoretical information on

measuring listening and speaking skills, practical content appeared to be limited. Additionally, it was found that teachers' participation in postgraduate professional development activities was relatively low. However, many teachers expressed a desire for further training in these areas, indicating a strong motivation to improve assessment literacy.

In conclusion, this study offers valuable insights into EFL teachers' assessment and evaluation practices related to oral language skills in Turkey. It highlights that assessment processes need enhancement both pedagogically and practically. For listening and speaking skills to become more prominent in teaching and for these skills to be assessed effectively, objectively, and sustainably, comprehensive interventions ranging from educational policies to teacher training are necessary.

Reflecting both as a teacher and researcher in that study, I actually realized that we need a change in mindset. Teachers need confidence, time collaboration, and they should acknowledge assessment as a part of learning, a learning process, rather than the result or an isolated event. As far as I observe and according to the conclusions, a group of them are aware and in favor of innovative ways of teaching and communicative assessment and authentic-based teaching, along with assessment; however, some of them are under the idea that implementation of assessment of speaking is time-consuming and tough with workload a lot. Actually, they are right on the issue of workload, but if we evaluate the process, I think this implementation has more advantages than disadvantages. I hope that this research contributes to the secondary school education field with views and experiences of Secondary School EFL teachers by emphasizing how crucial authentic, process-based assessment and fair assessment of listening and speaking as well as the possible steps and ideas to enhance conditions for speaking and listening assessment, as it is a timely issue that needs to be focused on well.

Quantitative Implications

Based on the findings of this study, important inferences can be made at various levels. These implications are noteworthy in terms of improving teaching practices and restructuring teacher training programs.

1. *Standardization of Assessment Processes Is Necessary:* The inconsistent scoring among teachers in speaking skills calls for the development of standard criteria, clear scoring guidelines, and uniform rubrics in this area. These tools make the assessment process more transparent and fairer.
2. *Encouraging alternative methods in listening and speaking assessments:* Alternative assessment methods can boost student participation and support a more comprehensive evaluation of skills, such as through portfolios, projects, and structured group work.

3. *Teacher Education Programs should be enhanced with practical content:* Measurement and evaluation courses in undergraduate teacher training should focus more on hands-on experience; real classroom scenarios, video analysis, and case-based training should be incorporated into the curriculum, especially for assessing oral skills.
4. *Professional Development Activities should be systematised:* seminars, workshops, and online training modules where teachers can enhance their listening and speaking skills should be provided regularly and accessibly.
5. *Systematic use of student feedback should be supported:* individual and structured feedback should be provided to students after both listening and speaking exams; ensure teachers are assisted with digital tools and feedback templates.
6. *Develop Real-Life-Based, Communication-Oriented Exam Formats:* Especially in speaking exams, create scenarios that reflect the communication situations students are likely to face. Such exams should be implemented widely in educational environments.
7. *Measurement and evaluation literacy should be built, and ongoing training opportunities should be offered to bridge the gap between teachers' theoretical understanding of assessment processes and their practical skills.*

In line with these conclusions, I believe that improvements at both the micro level (classroom practices) and the macro level (educational policies and teacher training) can support the effective measurement and evaluation of oral skills.

Like many of my colleagues, I would argue that there is a need for continuous reflection beyond peer exchange, and we require good practice examples to foster strong literacy in assessment. We urgently need structured and well-planned learning communities.

Qualitative Data Conclusions

This study demonstrated that integrating listening and speaking components into national English exams is generally beneficial for teachers, although it presents several practical challenges. Teachers saw the benefits of these assessments in boosting students' confidence, motivation, and communication skills. However, they also pointed out that structural and logistical issues, such as a lack of classroom equipment, limited time, and additional administrative work, make it challenging to carry out these assessments effectively.

The primary conclusion of this study is that a stronger infrastructure and improved teacher preparation are crucial for the success of assessment reforms. It is advised to shift towards an approach where oral skills are evaluated not by standalone, measurement-focused exams, but through continuous, classroom-based assessments. This change would not only alleviate performance anxiety but also enable a more genuine and meaningful evaluation of language use.

Qualitative Data Implications

These findings have important implications for educational policies, teaching practices, and future research. First, policymakers need to actively involve teachers and collaborate in piloting assessment reforms to ensure their relevance and practicality. Second, there is an urgent need for professional development programs that focus on designing and assessing oral performance-based tasks. Third, schools must be equipped with the necessary technological resources to ensure equal access to listening and speaking assessments. This study emphasizes the importance of integrating oral language into daily teaching and assessment practices. Expanding the use of task-based learning, peer interaction, and drama activities within the curriculum provides students with more opportunities to practice and assess their communication skills regularly.

Future research could expand on this study by exploring students' perspectives on oral assessment practices or comparing the long-term impacts of these practices on language skills and student motivation across various educational environments.

As a teacher-researcher working in a secondary school, I not only interpreted but also experienced many of the findings. Therefore, the implication suggests both scientific findings and my own reflections as a person who implemented both speaking and listening in the classrooms to my own students. As I have personally experienced, we need shared rubrics and common criteria that protect us from inconsistent results, even among every teacher in one school. As we do not have a common criteria, every teacher assesses individually, and results could change from class to class and student to student. In addition, I think all English teachers should be trained on gamification and digital tools that may be motivational for student involvement in speaking and listening assessment processes. Digital tools may also complement assessment by encouraging daily engagement of students, and teachers may provide formative feedback. Finally, this study, together with my colleagues, supports the idea that policy-level regulations are required. Classroom size is not appropriate for them. I agree with them because it is so difficult to benefit from the oral assessment process without manageable class sizes; we have limited time for every student to be engaged in speaking. As a practitioner, I also want to imply that authorities should consider the contextual issues that we face, and teachers' views from all over the country should be taken before the implementation process so that communication competence can be applied effectively and properly. They need to benefit from our experience, and the piloting process of these implementations should be carefully designed before being applied directly in schools.

I would like to conclude by saying that real assessment, I think, lies in the idea that assessment is a chance to learn, not an obligation, and if we all, teachers, policymakers, and students, can embrace this vision, transformation will start positively where we can listen and speak to each other well not through testing but by understanding each other's voices

REFERENCES

- Abugohar, M. A., Al-Hnifat, M. A., Al-Smadi, O. A., et al. (2019). English language speaking skill issues in an EMP context: Causes and solutions. *International Journal of English Linguistics*, 9(3), 211–225. <https://doi.org/10.5539/ijel.v9n3p211>
- Abuzied, A., & Nabag, W. (2023). Structured viva validity, reliability, and acceptability as an assessment tool in health professions education: A systematic review and meta-analysis. *BMC Medical Education*, 23(1). <https://doi.org/10.1186/s12909-023-04524-6>
- Akdoğan, N., & Toprak, E. (2024). The role of self-regulation and motivation in secondary school students' English self-efficacy. *Uluslararası Eğitim Araştırmacıları Dergisi*, 7(2), 82–104. <https://doi.org/10.52134/ueader.1435495>
- Akram, A., & Malik, A. (2010). Integration of language learning skills in second language acquisition. *International Journal of Arts and Sciences*, 3(14), 231–240.
- Alderson, J. C., & Huhta, A. (2005). The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22(3), 301–320.
- Alfa, R. R. (2020). The use of podcasts in improving students' speaking skill. *JELLT (Journal of English Language and Language Teaching)*, 4(1), 65–74. <https://doi.org/10.36597/jellt.v4i1.7692>
- Aşlıoğlu, B. (2009). Main listening barriers according to Turkish language student teachers. *Electronic Journal of Social Sciences*, 8(29), 45–63.
- Awad, M. J., & Al Adwan, M. A. (2024). Alternative assessment methods. In A. K. Abdallah, A. M. Alkaabi, & R. Al-Riyami (Eds.), *Cutting-edge innovations in teaching, leadership, technology, and assessment* (pp. 303–320). IGI Global. <https://doi.org/10.4018/978-1-3693-0880-6.ch020>
- Avedzi, J. (2025). Educational technology, teaching practices, and learning outcomes in Sub-Saharan Africa: Empirical insights from Ghana's basic education classrooms. <https://doi.org/10.21203/rs.3.rs-7330474/v1>
- Ayuanita, K. (2013). Assessing listening in the language classroom. *OKARA: Jurnal Bahasa dan Sastra*, 7(1).
- Bailey, K. D. (2008). *Methods of social research* (4th ed.). The Free Press.
- Bailey, K. M., & Savage, L. (1994). *New ways in teaching speaking* (2nd ed.).

- Barakat, A. (2023). The effects of digital drama-based instruction on developing receptive and expressive language among kindergarten children. *International Journal of Instruction*, 16(1), 103-118. <https://doi.org/10.29333/iji.2023.1616a>
- Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics*, 34(3), 304–324. <https://doi.org/10.1093/applin/ams046>
- Bashir, N. (2024). Teachers' perceptions of shifting from traditional exams to creative student responses: A study of the Bangladeshi education system. AUW.
- Bepari, N., & Imam, A. (2025). Enhancing English language skills through fun-focused teaching: An action research in Bangladeshi classrooms. *JOCW*, 9(1), 79-93. <https://doi.org/10.70771/jocw.163>
- Bilen, N., Koca, M., Antakyalı, A., Bağcı, E., Yakut, İ., Karaduman, M., Özat, M., Karakaya, E. H., Çetin, Ö. T., Sezgin, S., & Zerenli, A. K. (2022). Geçmişten günümüze ortaöğretimde, “Merkezi sınavlar”. *AA Agency*. <https://www.aa.com.tr/tr/pg/foto-galeri/gecmisten-gunumuze-ortaogretimde-merkezi-sinavlar/7>
- Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill Education (UK).
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). What makes listening difficult? Factors affecting second language listening comprehension.
- Borger, L. (2019). Assessing interactional skills in a paired speaking test. *Apples – Journal of Applied Language Studies*, 13(1), 151-174. <https://doi.org/10.17011/apples/urn.201903011694>
- Braney, B. T. (2011). *An examination of fourth grade teachers' assessment literacy and its relationship to students' reading achievement* (Unpublished master's thesis). University of Hartford.
- Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. In *Formative classroom assessment: Theory into practice* (Vol. 22, pp. 43–62).
- Brown, A., Iwashita, N., & McNamara, T. (2005). An examination of rater orientations and test-taker performance on English-for-academic-purposes speaking tasks. *ETS Research Report Series*, 2005(1), i–157. <https://doi.org/10.1002/j.2333-8504.2005.tb01982.x>

- Brown, H. D. (1990). M & Ms for language classrooms: Another look at motivation. In J. T. Alatis (Ed.), *Georgetown University round table on languages and linguistics* (pp. 383–393). Georgetown University Press.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.
- Buck, G. (1992). Listening comprehension: Construct validity and trait characteristics. *Language Learning*, 42(3), 313–357. <https://doi.org/10.1111/j.1467-1770.1992.tb01339.x>
- Buck, G. (2001). *Assessing listening*. Cambridge University Press.
- Bustari, A., Samad, I. A., & Achmad, D. (2017). The use of podcasts in improving students' speaking skill. *JELE (Journal of English Language and Education)*, 3(2), 97. <https://doi.org/10.26486/jele.v3i2.256>
- Caruso, M., Gadd Colombi, A., & Tebbit, S. (2017). Teaching how to listen: Blended learning for the development and assessment of listening skills in a second language. *Journal of University Teaching and Learning Practice*, 14(1). <https://doi.org/10.53761/1.14.1.7>
- Çelik, H., & Sur, E. (2025). Türkçe ve İngilizce öğretmenlerinin MEB tarafından uygulamaya konulan süreç odaklı ölçme-değerlendirmeye ilişkin görüşleri. *Bayburt Eğitim Fakültesi Dergisi*, 20(46), 636–653. <https://doi.org/10.35675/befdergi.1604619>
- Chen, Y. (2005). Barriers to acquiring listening strategies for EFL learners and their pedagogical implications. *Teaching English as a Second or Foreign Language*, 8(4).
- Cheng, H. (2004). A comparison of multiple-response formats for the assessment of listening. *Green*, 1975, 544–553.
- Cheung, A., & Wong, P. (2012). Factors affecting the implementation of curriculum reform in Hong Kong. *International Journal of Educational Management*, 26(1), 39–54. <https://doi.org/10.1108/09513541211194374>
- Choi, Y., Vo, S., & Ockey, G. (2023). Investigation into the factor structure of a local English placement test of oral communication. *International Journal of Applied Linguistics*, 34(1), 63–85. <https://doi.org/10.1111/ijal.12478>
- Cobbold, C., & Wright, L. (2021). Use of formative feedback to enhance summative performance. *Anatolian Journal of Education*, 6(1), 109–116. <https://doi.org/10.29333/aje.2021.619a>
- Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education* (5th ed.). Routledge. <https://doi.org/10.4324/9780203224342>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

- Colby-Kelly, C., & Turner, C. E. (2007). AFL research in the L2 classroom and evidence of usefulness: Taking formative assessment to the next level. *The Canadian Modern Language Review*, 64(1), 9–37. <https://doi.org/10.3138/cmlr.64.1.009>
- Cox, T., & Davies, R. (2012). Using automatic speech recognition technology with elicited oral response testing. *CALICO Journal*, 29(4), 601–618. <https://doi.org/10.11139/cj.29.4.601-618>
- Crossley, S., Clevinger, A., & Kim, Y. (2014). The role of lexical properties and cohesive devices in text integration and their effect on human ratings of speaking proficiency. *Language Assessment Quarterly*, 11(3), 250–270. <https://doi.org/10.1080/15434303.2014.926905>
- Daniel, A. K. (2011). Teachers and children: A classroom community of storytellers. In *Unlocking speaking and listening* (2nd ed., pp. 58–75). Routledge.
- Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. Teachers College Press.
- Davies, P., & Pearse, E. (2000). *Success in English teaching: A complete introduction to teaching English at secondary school level and above*. Oxford University Press.
- Davison, C. (2007). Views from the chalkface: English language school-based assessment in Hong Kong. *Language Assessment Quarterly*, 4(1), 37–68. <https://doi.org/10.1080/15434300701348359>
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393–415. <https://doi.org/10.1002/j.1545-7249.2009.tb00242.x>
- Dina Tsagari, & Banerjee, J. (2016). *Handbooks of inorganic synthesis*. Science Publication.
- Dewi, E., Hasanah, N., & Huda, M. (2020). Formative peer feedback on undergraduate students' speaking ability. *Al-Lisan*, 5(2), 116–129. <https://doi.org/10.30603/al.v6i2.1327>
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Duong, T. M. (2014). An investigation into effects of role-play in an EFL speaking course. *Global Journal of Foreign Language Teaching*, 4(2), 81–91.
- Elmankush, I. Y. (2017). Investigating the impact of including videos or still images in computer-based academic listening comprehension tests. University of York.
- Elshaw, N., Abdullah, A., & Rashid, S. (2017). Malaysian instructors' assessment beliefs in tertiary ESL classrooms. *International Journal of Education and Literacy Studies*, 5(2), 29–35. <https://doi.org/10.7575/aiac.ijels.v.5n.2p.29>
- Emiroğlu, S. (2013). Turkish teacher candidates' opinions about listening problems. *Adıyaman University Journal of Social Sciences*, 6(11), 269–307.

- Erickson, H. L. (2001). *Stirring the head, heart, and soul: Redefining curriculum and instruction*. Corwin Press.
- Fakhrurriana, R., & Herdina, G. F. (2024). Assessment test of receptive skills: Listening and reading comprehension of recount text using AI platform (Atomic Learning). *Proceeding International Conference on Religion, Science and Education*, 3, 187–192. <https://doi.org/10.18502/kss.v7i19.12480>
- Farhady, H. (2018). History of language testing and assessment. In *The TESOL encyclopedia of English language teaching* (pp. 1–7). Wiley. <https://doi.org/10.1002/9781118784235.eelt0343>
- Farooq, M. U. (2015). Creating a communicative language teaching environment for improving students' communicative competence at EFL/EAP university level. *International Education Studies*, 8(4). <https://doi.org/10.5539/ies.v8n4p179>
- Fauziah, Z., & Anisah, G. (2022). Challenges in assessing speaking proficiency in online learning during COVID-19 pandemic. *Challenges in Assessing Speaking Proficiency in Online Learning during COVID-19 Pandemic*, 13(1), 116-120.
- Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(1), 2. <https://doi.org/10.1186/s40468-019-0078-7>
- Flowerdew, J., & Miller, L. (2012). Assessing listening. In *The Cambridge guide to second language assessment* (pp. 225–233). Cambridge University Press.
- Frost, K., Elder, C., & Wigglesworth, G. (2011). Investigating the validity of an integrated listening-speaking task: A discourse-based analysis of test takers' oral performances. *Language Testing*, 29(3), 345–369. <https://doi.org/10.1177/0265532211424479>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Fulcher, G., & Davidson, F. (2013). *The Routledge handbook of language testing*. Routledge. <https://doi.org/10.4324/9780203181287>
- Garrison, C., & Ehringhaus, M. (2011). Formative and summative assessments in the classroom. *NMSA's Annual Conference and Exhibit*.
- Genesee, F., & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge University Press.
- Gilligan, M. E. (2007). Traditional versus alternative assessments: Which type do high school teachers perceive as most effective in the assessment of higher-order thinking skills? *Saint Louis University*.

- Ginther, A. (2012). In J. W. Oller Jr. (Ed.), *The encyclopedia of applied linguistics*. Wiley. <https://doi.org/10.1002/9781405198431.wbeal0882>
- Goh, C. C. M., & Vandergrift, L. (2021). *Teaching and learning second language listening: Metacognition in action*. Routledge. <https://doi.org/10.4324/9780429287749>
- Guetterman, T. C., & Fetters, M. D. (2018). Two methodological approaches to the integration of mixed methods and case study designs: A systematic review. *American Behavioral Scientist*, 62(7), 900–918. <https://doi.org/10.1177/0002764218772641>
- Hamouda, A. (2013). An investigation of listening comprehension problems encountered by Saudi students in the EL listening classroom. *International Journal of Academic Research in Progressive Education and Development*, 2(2), 113–155.
- Han, B., & Elçiçek, Z. (2021). Milli Eğitim Bakanlığı'nın 2023 eğitim vizyon belgesinin eğitim yönetimine ilişkin değişim alanlarının değerlendirilmesi. *Uluslararası Batı Karadeniz Sosyal ve Beşeri Bilimler Dergisi*, 5(1), 22–42. <https://doi.org/10.46452/baksoder.833404>
- Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4(2), 111–128.
- Hatipoğlu, Ç. (2017). Linguistics courses in pre-service foreign language teacher training programs and knowledge about language. *ELT Research Journal*, 6(1), 45–68.
- Haznedar, B. (2014). Cognitive and linguistic aspects of learning a second language in the early years. In *Early years second language education* (pp. 15–28). Routledge.
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Hedge, T. (2006). *Teaching and learning in the language classroom*. Oxford University Press.
- Hemsley, G., Holm, A., & Dodd, B. (2014). Identifying language difference versus disorder in bilingual children. *Speech, Language and Hearing*, 17(2), 101–115. <https://doi.org/10.1179/2050572813Y.0000000027>
- Huang, B., Bailey, A., Sass, D., & Chang, Y. (2020). An investigation of the validity of a speaking assessment for adolescent English language learners. *Language Testing*, 38(3), 401–428. <https://doi.org/10.1177/0265532220925731>
- Isakova, V., & Kaldybaev, S. (2025). The state of assessment organization in the educational process of educational institutions in Kyrgyzstan. *Bulletin of Science and Practice*, 11(8), 469–476. <https://doi.org/10.33619/2414-2948/117/65>

- Islam, W., Ahmad, S., & Islam, M. D. (2022). Investigating the problems faced by the university EFL learners in speaking English language. *International Journal of TESOL & Education*, 2(2), 47–65. <https://doi.org/10.54855/ijte.22223>
- Jamieson, J., & Poonpon, K. (2013). Developing analytic rating guides for TOEFL iBT's integrated speaking tasks. *ETS Research Report Series*, 2013(1). <https://doi.org/10.1002/j.2333-8504.2013.tb02320.x>
- Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2), 26–37.
- Jee, M. J. (2019). Foreign language anxiety and self-efficacy: Intermediate Korean as a foreign language learners. *Language Research*, 55(2), 431–456. <https://doi.org/10.30961/lr.2019.55.2.431>
- Joe, J., Kitchen, C., Chen, L., & Feng, G. (2015). A prototype public speaking skills assessment: An evaluation of human-scoring quality. *ETS Research Report Series*, 2015(2), 1–21. <https://doi.org/10.1002/ets2.12083>
- Kaldırım, A., & Degeç, H. (2017). Difficulties that students who learn Turkish as a foreign language encounter during listening skills. *Eğitim Bilimleri Araştırmaları Dergisi*, 7(1), 1–17.
- Kiomrs, R., Abdolmehdi, R., & Rashidi, N. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Teaching*, 4(1), 156–161.
- Koksal, D. (2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Online Journal of Distance Education*, 5(1).
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73(4), 440–464.
- Kusnierek, A. (2015). Developing students' speaking skills through role-play. *World Scientific News*, 7, 1–12.
- Lankara, V. V., Kittivajiro, P. S., & Thitavangso, P. V. (2024). Development of English listening and speaking skills for grade 10 students by using task-based learning at Myintkyina in Myanmar. *Journal of Asian Language Teaching and Learning*, 5(3).
- Lantolf, J. P., & Poehner, M. E. (2008). Dynamic assessment. In *Encyclopedia of language and education* (pp. 2406–2417). Springer.

- Lee, Y.-W. (2006). Dependability of scores for a new ESL speaking assessment consisting of integrated and independent tasks. *Language Testing*, 23(2), 131–166. <https://doi.org/10.1191/0265532206lt325oa>
- Leong, L. M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English speaking skill. *International Journal of English Education*.
- Lestari, L., & Asari, S. (2023). Investigating students' motivation in learning speaking through online learning. *Journal of Digital Learning and Distance Education*, 1(9), 338-345. <https://doi.org/10.56778/jdlde.v1i9.75>
- Long, D. R. (1989). Second language listening comprehension: A schema-theoretic perspective. *The Modern Language Journal*, 73(1), 32. <https://doi.org/10.2307/327265>
- Lucanus, A. (2017). Oral communication skills are important for students. Retrieved from The College Puzzle: <https://collegepuzzle.stanford.edu/oral-communication-skills-areimportant-for-students/>
- Lunar, L., Abril, N. D., Maningas, M. L., & Gloria, M. (2023). Honing science and language skills among student nurses through a study partner in biochemistry. *World Languages, Literature and Cultural Studies*, 2(2), 23–25. <https://doi.org/10.26480/wllcs.02.2023.23.25>
- Malone, M. E. (2017). Training in language assessment. In *Language testing and assessment* (pp. 225–239). Springer.
- Mart, Ç. T. (2020). Integrating listening and speaking skills to promote speech production and language development. *MEXTESOL Journal*, 44(2), 1–7.
- Maryslessor, A. O., Barasa, P. L., & Omulando, C. A. (2014). Challenges teachers face in the use of the communicative language teaching approach in the teaching listening and speaking lessons in Luger District, Kenya.
- Marzuki, D. (2024). The impact of explicit and implicit instruction on EFL learners' oral performance. *Australian Review of Applied Linguistics*, 48(3), 559-586. <https://doi.org/10.1075/aral.24029.ma>
- McNamara, T. F. (1997). "Interaction" in second language performance assessment: Whose performance? *Applied Linguistics*, 18(4), 446–466. <https://doi.org/10.1093/applin/18.4.446>
- Mertler, C. A. (2003). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?
- Mertler, C. A., & Campbell, C. (2005). Measuring teachers' knowledge & application of classroom assessment concepts: Development of the "Assessment Literacy Inventory".

- Milli Eğitim Bakanlığı Ölçme ve Değerlendirme Yönetmeliği. (2023). *Resmi Gazete*. <https://resmigazete.gov.tr/eskiler/2023/09/20230909-2.htm>
- Mirsanjari, Z. (2025). Enhancing assessment literacy in EAP instruction: The role of teacher development courses in overcoming systemic barriers. <https://doi.org/10.21203/rs.3.rs-6135411/v1>
- Miyane, S. (2020). Trends in interactional competence: Targeting engagement and explicit utterances. <https://doi.org/10.31219/osf.io/qhbvf>
- Mpanza, C., & Banja, M. (2025). An analysis of classroom assessment strategies utilized by secondary school teachers: A case of selected schools in Kabwe District, Zambia. *International Journal of Research and Innovation in Social Science*, 9(4), 1190-1205. <https://doi.org/10.47772/ijriss.2025.90400092>
- Msuya, P. (2025). Standardizing listening and speaking assessments to provide balanced coverage of essential skills. <https://doi.org/10.21203/rs.3.rs-7330474/v1>
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile Issues in Teachers' Professional Development*, 14(1), 143–158.
- Muthmainnah, A., Astri, Z., Inul, I., Jumardi, J., Tanasy, N., & Fachrunnisa, N. (2025). Teachers' effective strategies in reducing students' anxiety in speaking English. *GLENS*, 2(2), 83-91. <https://doi.org/10.61220/glens.v2i2.840>
- Nalani, A., Yoshikawa, H., & Way, N. (2024). Question asking in active listening scale for early adolescents: Behavioral measure development and initial validation. *Journal of Research on Adolescence*, 34(3), 745-758. <https://doi.org/10.1111/jora.12938>
- Newfields, T. (2006). Teacher development and assessment literacy. In *Authentic communication: Proceedings of the 5th Annual JALT Pan-SIG Conference* (pp. 48–73).
- Nihei, K. (2002). How to teach listening.
- Noroozi, M., & Taheri, S. (2021). The distinguishing characteristic of task-based language assessment. *Journal of Language Teaching and Research*, 12(5), 688–695. <https://doi.org/10.17507/jltr.1205.07>
- Nurharjanto, A. A., & Widyanoro, A. (2020). The effect of language learning strategy and technology toward students' writing skills. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, 4(2), 213–225. <https://doi.org/10.21831/jk.v4i2.34103>
- Ockey, G., Koyama, D., Setoguchi, E., & Sun, A. (2014). The extent to which TOEFL iBT speaking scores are associated with performance on oral language tasks and oral ability components for Japanese university students. *Language Testing*, 32(1), 39-62. <https://doi.org/10.1177/0265532214538014>

- Oktaviani, E., Ilham, I., Lukman, L., Olatunji, S., & Whitworth, N. (2024). The effect of show and tell, role-play and storytelling on speaking skills: Meta-analysis. *Journal of Languages and Language Teaching*, 12(2), 846-859. <https://doi.org/10.33394/joltt.v12i2.10868>
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Addison-Wesley Publishing Company.
- Ölmezer-Öztürk, E., & Aydin, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of Language Assessment Knowledge Scale (LAKS). *Language Testing in Asia*, 8, 1–15.
- Oo, T. Z., Habók, A., & Józsa, K. (2023). Qualifying method-centered teaching approaches through the reflective teaching model for reading comprehension. *Education Sciences*, 13(5), 473. <https://doi.org/10.3390/educsci13050473>
- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3(1), 53–66.
- Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25–44.
- Palmer, E. (2014). *Teaching the core skills of listening and speaking*. ASCD.
- Paneerselvam, A., & Mohamad, M. (2019). Learners' challenges and English educators' approaches in teaching speaking skills in an ESL classroom: A literature review. *Creative Education*, 10(13), 3299–3305. <https://doi.org/10.4236/ce.2019.1013253>
- Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, 8, 1217167. <https://doi.org/10.3389/feduc.2023.1217167>
- Prastikawati, F., Mujiyanto, E., Saleh, J., & WuliFitriati, S. M. (2022). Pre-service EFL teachers' conceptions of assessment during their teaching practicum. *KnE Social Sciences*. <https://doi.org/10.18502/kss.v7i19.12480>
- Peltekov, P. (2022). Investigating the validity of the IELTS listening test. *GJELT*, 2(1), 9-15. <https://doi.org/10.20448/gjelt.v2i1.4128>
- Prahl, A., & Schuele, C. (2022). A pilot study assessing listening comprehension and reading comprehension in children with Down syndrome: Construct validity from a multi-method perspective. *Frontiers in Psychology*, 13, Article 905273. <https://doi.org/10.3389/fpsyg.2022.905273>
- Qayoom, N., Saleem, M., & Mansoor, M. (2021). Objectives of teaching English skills and their assessment in Indian schools: A study on the divergence. *Journal of Language and Linguistic Studies*, 17(2), 755–766. <https://search.informit.org/doi/10.3316/informit.216174064973954>

- Rahmawati, Y., & Ertin, E. (2014). Developing assessment for speaking. *IJEE (Indonesian Journal of English Education)*, 1(2), 199–210. <https://doi.org/10.15408/ijee.v1i2.1345>
- Rajendran, T., & Md Yunus, M. (2021). A systematic literature review on the use of mobile-assisted language learning (MALL) for enhancing speaking skills among ESL and EFL learners. *International Journal of Academic Research in Progressive Education and Development*, 10(1), 586–609. <https://doi.org/10.6007/ijarped/v10-i1/8939>
- Rajput, N. H., Chandio, M. Y., Bhutto, R. N., & Kanhio, Z. H. (2020). Headteachers' perspectives about the English language teaching at lower secondary level in Shaheed Benazirabad, Pakistan. *Journal of Literature, Languages and Linguistics*, 71, 55–61. <https://doi.org/10.7176/JLLL/71-04>
- Rao, P. S. (2019). The importance of speaking skills in English classrooms. *Alford Council of International English & Literature Journal (ACIELJ)*, 2(2), 6–18.
- Reeves, T. C. (2000). Alternative assessment approaches for online learning environments in higher education. *Journal of Educational Computing Research*, 23(1), 101–111. <https://doi.org/10.2190/GYMQ-78FA-WMTX-J06C>
- Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.
- Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of AI technologies in high-stakes assessment. *London Review of Education*, 19(1). <https://doi.org/10.14324/LRE.19.1.09>
- Rost, M. (2006). Areas of research that influence L2 listening instruction. *Current Trends in the Development and Teaching of the Four Language Skills*, 47, 73.
- Rubin, J. (1990). Improving foreign language listening comprehension. *Georgetown University Round Table on Languages and Linguistics*, 309–316.
- Rubin, R. B. (2009). Assessing speaking and listening competence at the college level: The communication competency assessment instrument. *Communication Education*, 31(1), 19–32. <https://doi.org/10.1080/03634528209384656>
- Sabina, Z. (2018). The importance of teaching listening and speaking skills. *World Science*, 7(34), 52–55.
- Saeed, K., Ismail, S., & Eng, L. (2019). Malaysian speaking proficiency assessment effectiveness for undergraduates suffering from minimal descriptors. *International Journal of Instruction*, 12(1), 1059–1076. <https://doi.org/10.29333/iji.2019.12168a>

- Sahyoni, S., & Zaim, M. (2017). Authentic assessment of speaking skill for grade I junior high school. *Komposisi: Jurnal Pendidikan Bahasa, Sastra, dan Seni*, 18(1), 15-27. <https://doi.org/10.24036/komposisi.v18i1.6509>
- Saidalvi, A., & Samad, A. (2019). Online peer motivational feedback in a public speaking course. *GEMA Online Journal of Language Studies*, 19(4), 258-277. <https://doi.org/10.17576/gema-2019-1904-14>
- Salekhova, L., Grigorieva, K., Aristova, N., & Gerasimova, A. (2019). Assessing multilingual competence basing on holistic approach. In *13th International Technology, Education and Development Conference* (pp. 2960–2965). <https://doi.org/10.21125/inted.2019.0782>
- Samad, I. A., & Adnan, Z. (2015). Using a genre-based approach to prepare undergraduate students for an English thesis defence examination: An experimental study to address the ‘pedagogical controversy’. *Linguistik Indonesia*, 35(1), 75–93.
- Sanavi, R. V., & Mohammadi, M. (2018). Listening assessment. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). <https://doi.org/10.1002/9781118784235.eelt0806>
- Saptiany, S., & Prabowo, B. (2024). Speaking proficiency among English-specific-purpose students: A literature review on assessment and pedagogical approaches. *Literacy International Scientific Journals of Social Education Humanities*, 3(1), 36-48. <https://doi.org/10.56910/literacy.v3i1.1392>
- Sarwar, M., Alam, M., Hussain, S., Shah, A. A., & Jabeen, M. (2014). Assessing English speaking skills of prospective teachers at entry and graduation level in teacher education program. *Language Testing in Asia*, 4(1), 5.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.
- Shalash, M. J. (2024). Advancing language assessment and evaluation: Innovative strategies for specialized English teaching. *Interdisciplinary Journal Papier Human Review*, 5(1), 1–12. <https://doi.org/10.47667/ijphr.v5i1.297>
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263-300. <https://doi.org/10.1191/1362168804lr146oa>
- Sheetz, D., Coldwell, S., & Coombe, C. (2018). Speaking assessment. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–8). <https://doi.org/10.1002/9781118784235.eelt0338>
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.

- Siddiqui, N., Khan, R., Afifa, S., & Islam, Z. (2022). Perception and validity of online formative oral assessment in surgical trainees: A cross-sectional study. *Cureus*, 14(7), e24938. <https://doi.org/10.7759/cureus.24938>
- Slamet, J., & Mukminatien, N. (2024). Language education and acquisition research network. *Language Education and Acquisition Research Network*, 17(1).
- Srikaew, D., Tangdhanakanond, K., & Kanjanawasee, S. (2015). Development of an English speaking skill assessment model for grade 6 students by using portfolio. *Procedia - Social and Behavioral Sciences*, 191, 764–768. <https://doi.org/10.1016/j.sbspro.2015.04.715>
- Stake, R. E. (1995). *The art of case study research*. Sage Publications.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534–539.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5–15. <https://doi.org/10.1111/j.1745-3992.2001.tb00065.x>
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758–765. <https://doi.org/10.1177/003172170208301010>
- Stošić, L., & Malyuga, E. N. (2024). Application of artificial intelligence in language skills testing. *Anglisticum. Journal of the Association-Institute for English Language and American Studies*, 13(1), 22–34. <https://doi.org/10.58885/ijllis.v13i1.22ls>
- Swain, M., Huang, L., Barkaoui, K., Brooks, L., & Lapkin, S. (2009). The PEaking section of the TOEFL iBT™ (SSTIBT): Test-takers' reported strategic behaviors. In *ETS Research Report Series* (No. 2). <https://doi.org/10.1002/j.2333-8504.2009.tb02187.x>
- Tleuov, A. (2025). Teaching speaking in Kazakhstani EFL classrooms: Negotiating teacher beliefs and assessment constraints. *IAFOR Journal of Education*, 13(1), 11-41. <https://doi.org/10.22492/ije.13.1.01>
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63.
- Turner, J. (1998). Assessing speaking. *Annual Review of Applied Linguistics*, 18, 192–207. <https://doi.org/10.1017/S0267190500003548>
- Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge.
- Underwood, M., Kenworthy, J., & Rost, M. (1989). *Teaching listening* (C. 117). Longman.

- Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies.
- Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *Modern Language Journal*, 90(1), 6–18. <https://doi.org/10.1111/j.1540-4781.2006.00381.x>
- Vandergrift, L. (2011). Second language listening: Presage, process, product, and pedagogy. In *Handbook of research in second language teaching and learning* (1st ed., pp. 455–471). Routledge.
- Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, 65(2), 390–416. <https://doi.org/10.1111/lang.12105>
- Vandergrift, L., & Goh, C. (2012). Teaching and learning second language listening: Metacognition in action. *The CATESOL Journal*, 27(2), 303–311.
- Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, 60(2), 470–497. <https://doi.org/10.1111/j.1467-9922.2009.00559.x>
- Verschuren, P. (2003). Case study as a research strategy: Some ambiguities and opportunities. *International Journal of Social Research Methodology*, 6(2), 121–139. <https://doi.org/10.1080/13645570110106154>
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749–770.
- Voss, B. (1979). Hesitation phenomena as sources of perceptual errors for non-native speakers. *Language and Speech*, 22(2), 129–144.
- White, E. (2009). Are you assessment literate? Some fundamental questions regarding effective classroom-based assessment. *OnCUE Journal*, 3(1), 3–25.
- Woottipong, K. (2014). Effect of using video materials in the teaching of listening skills for university students. *International Journal of Linguistics*, 6(4), 200.
- Xu, J. (2018). Measuring “spoken collocational competence” in communicative speaking assessment. *Language Assessment Quarterly*, 15(3), 255–272. <https://doi.org/10.1080/15434303.2018.1482900>
- Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Studies in Language Assessment*, 6(1), 133–158.

- Yang, Y. I. J. (2014). Is speaking fluency strand necessary for the college students to develop in the EFL class? *Theory and Practice in Language Studies*, 4(2), 225–231. <https://doi.org/10.4304/tpls.4.2.225-231>
- Yastıbaşı, A. E., & Takkaç, M. (2018). Understanding language assessment literacy: Developing language assessments. *Journal of Language and Linguistic Studies*, 14(1), 178–193.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications, Inc.
- Yıldırım, A., & Şimşek, H. (2011). *Sosyal bilimlerde nitel araştırma yöntemleri* (8. basım). Seçkin Yayıncılık.
- Yıldız, M., Kabaran, H., & Yaşar Sağlık, Z. (2024). Ortaöğretime geçiş sınavlarının (OKS, SBS, TEOG, LGS) kelime sayıları ve sıklıkları bakımından incelenmesi. *Türk Eğitim Bilimleri Dergisi*, 22(2), 919–943. <https://doi.org/10.37217/tebd.1336962>
- Yu, M. H., Reynolds, B. L., & Ding, C. (2021). Listening and speaking for real-world communication: What teachers do and what students learn from classroom assessments. *SAGE Open*, 11(2), 21582440211009164. <https://doi.org/10.1177/21582440211009163>
- Yustina, L., Besral, B., & Hasnawati, H. (2021). Classroom assessment for EFL learning to speak at junior high school. *Al-Ta'lim*, 28(2), 134-144. <https://doi.org/10.15548/jt.v28i2.702>
- Yuzar, E. (2020). Incorporating communicative competence in assessment and English language teaching in multilingual settings. *Journal of Research and Innovation in Language*, 2(1), 8–13. <https://doi.org/10.31849/reila.v2i1.3864>
- Zaim, M., & Arsyad, S. (2020). Authentic assessment for speaking skills: Problem and solution for English secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604.
- Zedelius, C., Mills, C., & Schooler, J. (2018). Beyond subjective judgments: Predicting evaluations of creative writing from computational linguistic features. *Behavior Research Methods*, 51(2), 879-894. <https://doi.org/10.3758/s13428-018-1137-1>
- Zhu, X. (2011). Study on the influence of voice on listening comprehension. *2011 IEEE 3rd International Conference on Communication Software and Networks*, 749–753. <https://doi.org/10.1109/ICCSN.2011.6014198>
- Zuber, C., & Conzelmann, A. (2019). Achievement-motivated behavior in individual sports (AMBIS-I): Coach rating scale. *German Journal of Exercise and Sport Research*, 49(4), 410-423. <https://doi.org/10.1007/s12662-019-00600-6>



APPENDICES

Appendix-1: Questionnaire

ANKET KATILIMCISININ DİKKATİNE

- 1. Araştırmanın Amacı:** Bu çalışma, İngilizce öğretmenlerinin konuşma ve dinleme sınavlarıyla ilgili görüşlerini ve deneyimlerini derinlemesine incelemeyi hedeflemektedir. Bu bağlamda, öğretmenlerin sınavların etkinliği, öğrenci başarısı üzerindeki etkileri ve sınav süreçlerinin iyileştirilmesi konularındaki görüşlerini anlamak amaçlanmaktadır.
- 2. Katılımın Gönüllülüğü ve Geri Çekme Hakkı:** Katılım tamamen gönüllülük esasına dayanmaktadır. Katılımcılar, herhangi bir neden göstermeksizin ve herhangi bir zamanda araştırmadan çekilebilme hakkına sahiptirler. Bu durum, katılımcıların araştırmaya katılımının kişisel tercihlerine bağlı olduğunu ve katılmamalarının veya geri çekilmelerinin herhangi bir olumsuz sonuç doğurmayacağını garanti altına almaktadır.
- 3. Araştırmanın Yararlanacakları:** Araştırmanın sonuçları, özellikle İngilizce öğretmenleri, eğitim kurumları ve eğitim politikası oluşturucuları için değerli olacaktır. Bu bulguların, eğitimde sınav süreçlerinin daha etkili ve adil hale getirilmesi için nasıl kullanılabileceği üzerine önemli bir katkı sağlaması beklenmektedir.
- 4. Bilgi Yayınlanacak Yerler:** Araştırma bulguları, uluslararası akademik dergilerde, eğitim konferanslarında sunulacak ve üniversitenin dijital arşivinde yayınlanacaktır. Bu şekilde, araştırma sonuçlarının geniş bir kitleye ulaşması ve eğitim alanında bilimsel tartışmalara katkıda bulunması hedeflenmektedir.
- 5. Veri Erişimi ve Gizlilik:** Araştırmada kullanılan verilere sadece araştırma ekibi erişebilecek ve veriler üçüncü şahıslarla paylaşılmayacaktır. Katılımcıların gizliliği ön planda tutulacak, veriler anonim olarak saklanacak ve kişisel bilgilerin korunması sağlanacaktır.
- 6. Ek Bilgi Talebi ve İletişim Bilgileri:** Katılımcılar, araştırma hakkında daha fazla bilgi almak veya sorularını yanıtlamak için araştırma ekibiyle iletişime geçebileceklerdir. Bu bağlamda, araştırma ekibinin sorumlu kişinin adı ve iletişim bilgileri anket belgesinde yer alacaktır.

Appendix-1: (Continue)

KISIM-I DEMOGRAFİK BİLGİLER

Bu bölümde, anketin sonuçlarını daha iyi analiz edebilmek amacıyla demografik bilgilere yer veriyoruz. Verdiğiniz yanıtlar sadece istatistiksel amaçlarla kullanılacaktır.

Cinsiyetiniz?

- a) Kadın b) Erkek

Yaşınız?

NOT: Lütfen **RAKAM** ile yazınız.

Mezun olduğunuz bölüm hangisi?

- a) İngilizce Öğretmenliği b) İngiliz Dili “ve Edebiyatı” c) Amerikan Dili ve Edebiyatı
d) Mütercim Tercümanlık e) Diğer (Lütfen Belirtiniz.)

Eğitim dereceniz?

- a) Lisans b) Yüksek Lisans c) Doktora

Ne kadar süredir öğretmenlik yapıyorsunuz?

- a) 1 yıldan az b) 1-3 yıl c) 4-7 yıl d) 8-11 yıl
e) 12-15 yıl f) 16-19 yıl g) 20 yıldan fazla

KISIM-II: DİNLEME BECERİLERİNİN DEĞERLENDİRİLMESİ

Bu bölümde dinleme becerilerinin değerlendirmeye dahil edilmesine ilişkin düşünceleriniz ile alakalı sorular yer almaktadır.

A: Dinleme Sınavlarının Uygulanması ve Değerlendirilmesi

Bu soru grubu, İngilizce öğretmenlerinin dinleme sınavlarını nasıl uyguladığını, hangi yöntemleri ve araçları tercih ettiğini kapsamaktadır.

S1. Dinleme becerilerini değerlendiren sınavları ne sıklıkla kullanıyorsunuz?

- a) Haftalık b) Aylık c) Dönemlik
d) Yıllık e) Hiç kullanmıyorum.

S2. Dinleme becerilerini değerlendirirken hangi sınav formatını tercih ediyorsunuz?

- a) Çoktan seçmeli sorular b) Boşluk doldurma
c) Açık uçlu sorular d) Doğru/Yanlış

S3. Dinleme becerilerini değerlendirmede hangi faktörlerin önemli olduğunu düşünüyorsunuz? (NOT: Birdenfala seçenek işaretleyebilirsiniz.)

- a) Görev açıklığı b) Materyal kalitesi c) Zaman yönetimi
d) Öğrenci motivasyonu e) Becerilerin ölçümü f) Diğer

Appendix-1: (Continue)

S4. Dinleme becerilerini değerlendirmede etkin faktörlerin zorluk derecelerini belirtiniz?

1: Çok Kolay, 5: Çok Zor	1	2	3	4	5
a) Zaman yönetimi	☐	☐	☐	☐	☐
b) Materyal uyumluluğu	☐	☐	☐	☐	☐
c) Öğrenci katılımı	☐	☐	☐	☐	☐
e) Değerlendirme objektifliği	☐	☐	☐	☐	☐

S5. Öğrencilerin **dinleme becerilerini değerlendirmek için hangi ölçme araçları kullanılmalı?**
(*NOT: Birden fazla seçenek işaretleyebilirsiniz.*)

- a) Öğrenci performansı b) Sınavlar c) Ödevler
d) Portfolio değerlendirmesi e) Diğer

S6. En son **dinleme sınavında hangi ölçme araçlarını kullandınız?**

Lütfen belirtiniz: _____.

S7. Sizce dinleme becerilerini değerlendirmek için aşağıda verilenlerden hangisi(leri) kullanılmalıdır? (NOT: Birden fazla seçenek işaretleyebilirsiniz.**)**

- a) Günlük konuşmalar/Diyaloglar b) Sunumlar c) Röportajlar
d) Radyo programlar e) Müzik parçaları f) Diğer

S8. Sizce **dinleme becerilerini değerlendirmede en etkili sınav formatı hangisidir?**

- a) Dinleme ve soru-cevap b) Dinleme ve hikâye anlatma e) Diğer
d) Dinleme ve diyalog tamamlama e) Dinleme ve anlama soruları

S9. Yaptığınız **dinleme sınavlarında hangi stratejilerin daha etkili olduğunu düşünüyorsunuz?**

- a) Örnek materyal kullanma
b) Çeşitli sınav formatları (çoktan seçmeli, True/False vb.) kullanma
c) İçerik ve görev çeşitliliği sağlama
d) Öğrenci geri bildirimini düzenli olarak sağlama
e) Diğer

S10. Sizce başka bir öğretmen sizin yaptığınız **dinleme sınavını değerlendiği takdirde benzer ya da aynı puanları verir miydi?**

- a) Daha çok puan verirdi. b) Benzer puanlar verirdi.
c) Aynı puanı verirdi. d) Daha düşük puan verirdi.

Appendix-1: (Continue)

B: Dinleme Sınavlarının Güvenilirliği ve Geçerliliği

Bu soru grubu, öğretmenlerin uyguladığı sınavların güvenilirliği, tutarlılığı ve değerlendirme süreçlerinin doğruluğu ile alakalıdır.

S11. Sizce yaptığınız dinleme sınavının sonuçları gerçek dünyadaki konuşma performansını ne kadar yordamaktadır?

- a) Çok yüksek düzeyde yordadığını düşünüyorum.
- b) Yüksek **düzeyde** yordadığını düşünüyorum.
- c) Orta düzeyde yordamaktadır.
- d) Düşük düzeyde yordamaktadır.
- e) Hiç yordamamaktadır.

S12. Sizce yaptığınız dinleme sınavları genel olarak öğretmen dostu muydu? (Uygulanması ve değerlendirmesi kolay mı?)

- a) Kesinlikle öğretmen dostuydu.
- b) Öğretmen dostuydu.
- c) Ne öğretmen dostu ne değil.
- d) Öğretmen dostu değildi.
- e) Kesinlikle öğretmen dostu değildi.

S13. Dinleme sınavlarınızın yeterince güvenilir (tutarlı) sonuçlar verdiğini düşünüyor musunuz?

- a) Kesinlikle evet
- b) Evet, çoğunlukla
- c) Kararsızım
- d) Hayır, pek değil
- e) Kesinlikle hayır

S14. Yaptığınız dinleme sınavları yeterince geçerli miydi (amaçlananı ölçmüş müydü)?

- a) Kesinlikle evet
- b) Evet, çoğunlukla
- c) Kararsızım
- d) Hayır, pek değil
- e) Kesinlikle hayır

S15. Dinleme sınavlarının güvenilirliğini, geçerliliğini veya genel kalitesini artırmak için herhangi bir öneriniz var mı?

- a) Evet
- b) Hayır

S16. Dinleme sınavlarının güvenilirliğini, geçerliliğini veya genel kalitesini artırmak için önerileriniz nelerdir?

Lütfen kısaca yanıtlayınız: _____.

C: Dinleme Sınavlarının Öğretme ve Öğrenmeye Etkisi

Bu soru grubu, öğretmenlerin sınavların öğrencilere sağladığı katkıyı ve öğretim süreçlerine etkisini nasıl değerlendirdiğini ele alır.

S17. Genel olarak **dinleme** becerisi uygulama sınavından memnuniyetiniz nedir?

- a) Çok memnunum
- b) Biraz memnunum
- c) Ne memnunum ne değilim
- d) Pek memnun değilim
- e) Memnun değilim

Appendix-1: (Continue)

S18. Dinleme sınav sonuçlarını kullanarak öğrencilere destek sağladınız mı?

a) Evet

b) Hayır

S19. Dinleme sınav sonuçlarını kullanarak öğrencilere ne tür bir destek sağladınız? (**NOT: Birden fazla seçenek işaretleyebilirsiniz?**)

☐ Ek ders materyalleri

☐ Bireysel danışmanlık

☐ Grup çalışmaları

☐ Geri bildirim oturumları

S20. Dinleme sınavları, öğrencileri becerilerini geliştirmeleri için ne ölçüde motive etti?

a) Çok yüksek düzeyde motive etti.

b) Yüksek düzeyde motive etti.

c) Orta düzeyde motive etti.

d) Düşük düzeyde motive etti.

e) Hiç motive etmedi.

21. Dinleme sınavlarının öğretme ve öğrenmeyi ne kadar etkilediğini düşünüyorsunuz?

a) Çok olumlu etkiliyor

b) Olumlu etkiliyor

c) Ne olumlu ne olumsuz

d) Olumsuz etkiliyor

e) Çok olumsuz etkiliyor

S22. Yaptığınız dinleme sınavlarının öğrencilerin dil gelişimine yeterince katkısı olduğunu düşünüyor musunuz?

a) Kesinlikle evet

b) Evet, çoğunlukla

c) Kararsızım

d) Hayır, pek değil

e) Kesinlikle hayır

D: Öğretmenlerin Geri Bildirim Kullanımı

Bu soru grubu, öğretmenlerin sınav sonuçlarını nasıl değerlendirdiğini ve öğrencilere geri bildirim sağlama süreçlerini inceler.

S23. Dinleme sınavı değerlendirme sürecinde öğrenci geri bildirimini kullanıyor musunuz?

a) Evet

b) Hayır

S24. Dinleme sınavı değerlendirme sürecinde öğrenci geri bildirimlerini nasıl kullanıyorsunuz? (**NOT: Birden fazla seçenek işaretleyebilirsiniz?**)

☐ İyileştirme önerileri sunarak

☐ Öğrencilerin güçlü ve zayıf yönlerini vurgulayarak

☐ İlerleme raporları hazırlayarak

☐ Öğrenci-öğretmen görüşmeleri yaparak

☐ Diğer

S25. Dinleme sınavı değerlendirme sürecinde öğrenci geribildirimini yararlı olacağını düşünüyor musunuz?

a) Evet

b) Hayır

Appendix-1: (Continue)

S26. Dinleme sınavı değerlendirme sürecinde öğrenci geri bildirimini nasıl kullanıyorsunuz?

Lütfen kısaca yanıtlayınız: _____.

S27. Öğrencilerin **dinleme** becerilerini geliştirmek için onlara herhangi bir geribildirim sağlıyor musunuz?

a) Evet

b) Hayır

S28. Öğrencilerin **dinleme** becerilerini geliştirmek için onlara hangi geribildirmiyi sağlıyor musunuz?

- a) Gramer ve telaffuz hatalarını düzeltme
- b) İçerik ve anlamı değerlendirme
- c) Konuşma akıcılığına ve anlatım becerisine odaklanma
- d) Öğrencileri güçlü yönlerine teşvik etme
- e) Diğer

S29. Öğrencilerin **dinleme** becerilerini değerlendirirken hangi kriterleri kullanıyorsunuz? (**NOT: Birden fazla seçenek işaretleyebilirsiniz?**)

- | | |
|--|--|
| ☐ Anlama düzeyi | ☐ Sözcüklerin telaffuzunu anlama |
| ☐ Ana fikri/konuyu kavrama | ☐ Tonlama ve vurgulamayı anlama |
| ☐ Detayları anlama | ☐ Dinleme sonrası çıkarım yapabilme |
| ☐ Kelime dağarcığı zenginliği | ☐ Diğer |
| ☐ Anahtar kavramları/kelimeleri tanım | |

E: Öğretmenlerin Mesleki Gelişim ve Eğitimi Geçmişi

Bu soru grubu, öğretmenlerin dinleme becerilerinin ölçülmesine yönelik aldıkları eğitimi ve mesleki gelişim süreçlerini kapsamaktadır.

S30. Lisans eğitiminizde **dinleme** becerilerinin ölçülmesine yönelik yeterli bir eğitim aldınız mı?

a) Evet

b) Hayır

S31. Lisans eğitiminizde **dinleme** becerilerinin ölçülmesine yönelik aldığınız bu eğitimi nasıl değerlendiriyorsunuz?

- | | | |
|-------------|---------------------|---------|
| a) Çok iyi | b) İyi | c) Orta |
| d) Yetersiz | e) Oldukça yetersiz | |

S32. Lisans sonrasında **dinleme** becerilerinin ölçülmesine yönelik herhangi bir mesleki etkinliğe (konferans, seminer, webinar vb.) katıldınız?

a) Evet

b) Hayır

S33. **Dinleme** becerilerinin ölçülmesine yönelik bir eğitime katılmak ister miydiniz? Böylesi bir etkinliği ne derece yararlı olacağını düşünüyorsunuz?

- a) Evet, katılmak isterim ve çok yararlı olacağını düşünüyorum
- b) Evet, katılmak isterim ve yararlı olacağını düşünüyorum.
- c) Kararsızım.
- d) Hayır, katılmak istemem ama yararlı olacağını düşünüyorum.
- e) Hayır, katılmak istemem ve yararlı olacağını düşünmüyorum.

Appendix-1: (Continue)

S34. Şu an **dinleme** becerilerinin ölçülmesine yönelik aşağıdaki mesleki etkinliklerinden hangisi sizin için daha verimli olur? (**NOT:** Birden fazla seçenek işaretleyebilirsiniz?)

- | | |
|---|---|
| ☐ Konferanslar | ☐ Sunumlar |
| ☐ Seminerler | ☐ Münazaralar |
| ☐ Webinarlar | ☐ Eğitim Yarışmaları |
| ☐ Panel Tartışmaları | ☐ Çevrimiçi Dersler |

KISIM-III: KONUŞMA BECERİLERİNİN DEĞERLENDİRİLMESİ

Bu bölümde konuşma becerilerinin değerlendirmeye dahil edilmesine ilişkin düşünceleriniz ile alakalı sorular yer almaktadır.

A: Konuşma Sınavlarının Uygulanması ve Değerlendirilmesi

Bu soru grubu, İngilizce öğretmenlerinin dinleme sınavlarını nasıl uyguladığını, hangi yöntemleri ve araçları tercih ettiğini kapsamaktadır.

S35. Konuşma becerilerini değerlendiren sınavları ne sıklıkla kullanıyorsunuz?

- | | | |
|-------------|----------------------|-------------|
| a) Haftalık | b) Aylık | c) Dönemlik |
| d) Yıllık | e) Hiç kullanmıyorm. | |

S36. Konuşma becerilerini değerlendirirken hangi sınav formatını tercih ediyorsunuz?

- | | | |
|-----------------------|-------------|-------------|
| a) Rol yapma | b) Tartışma | c) Sunumlar |
| d) Drama etkinlikleri | e) Diğer | |

S37. Konuşma becerilerini değerlendirmede hangi faktörlerin önemli olduğunu düşünüyorsunuz? (**NOT:** Birden fazla seçenek işaretleyebilirsiniz.)

- | | | |
|------------------------|-----------------------|-------------------|
| a) Görev açıklığı | b) Materyal kalitesi | c) Zaman yönetimi |
| d) Öğrenci motivasyonu | e) Becerilerin ölçümü | f) Diğer |

S38. Konuşma becerilerini değerlendirmede etkin faktörlerin zorluk derecelerini belirtiniz?

1: Çok Kolay, 5: Çok Zor	1	2	3	4	5
a) Zaman yönetimi	☐	☐	☐	☐	☐
b) Materyal uyumluluğu	☐	☐	☐	☐	☐
c) Öğrenci katılımı	☐	☐	☐	☐	☐
e) Değerlendirme objektifliği	☐	☐	☐	☐	☐

S39. Öğrencilerin **konuşma** becerilerini değerlendirmek için hangi ölçme araçları kullanılmalı? (**NOT:** Birden fazla seçenek işaretleyebilirsiniz.)

- | | | |
|------------------------------|-------------|-------------------|
| a) Öğrenci performansı | b) Sınavlar | c) Proje Ödevleri |
| d) Portfolio değerlendirmesi | e) Diğer | |

Appendix-1: (Continue)

S40. En son **konuşma** sınavında hangi ölçme araçlarını kullandınız?

Lütfen belirtiniz: _____.

S41. Sizce **konuşma** becerilerini değerlendirmek için aşağıda verilenlerden hangisi(leri) kullanılmalıdır? (**NOT: Birden fazla seçenek işaretleyebilirsiniz.**)

- | | | |
|---------------------------------|----------------|----------------|
| a) Günlük konuşmalar/Diyaloglar | b) Sunumlar | c) Röportajlar |
| d) Tartışmalar | e) Münazaralar | f) Diğer |

S42. Sizce **konuşma** becerilerini değerlendirmede en etkili sınav formatı hangisidir?

- | | | |
|---------------------------------|-------------------------------|----------|
| a) Konuşma ve soru-cevap | b) Konuşma ve hikâye anlatma | e) Diğer |
| c) Konuşma ve diyalog tamamlama | d) Konuşma ve anlama soruları | |

S43. Yaptığınız **konuşma** sınavlarında hangi stratejilerin daha etkili olduğunu düşünüyorsunuz?

- a) Örnek materyal kullanma
- b) Çeşitli sınav formatları (çoktan seçmeli, True/False vb.) kullanma
- c) İçerik ve görev çeşitliliği sağlama
- d) Öğrenci geri bildirimini düzenli olarak sağlama
- e) Diğer

S44. Sizce başka bir öğretmen sizin yaptığınız **konuşma** sınavını değerlendiği takdirde benzer ya da aynı puanları verir miydi?

- | | |
|---------------------------|-----------------------------|
| a) Daha çok puan verirdi. | b) Benzer puanlar verirdi. |
| c) Aynı puanı verirdi. | d) Daha düşük puan verirdi. |

B: Konuşma Sınavlarının Güvenilirliği ve Geçerliliği

Bu soru grubu, öğretmenlerin uyguladığı sınavların güvenilirliği, tutarlılığı ve değerlendirme süreçlerinin doğruluğu ile alakalıdır.

S45. Sizce yaptığınız **konuşma** sınavının sonuçları gerçek dünyadaki konuşma performansını ne kadar yordamaktadır?

- a) Çok yüksek düzeyde yordadığını düşünüyorum.
- b) Yüksek **düzeyde** yordadığını düşünüyorum.
- c) Orta düzeyde yordamaktadır.
- d) Düşük düzeyde yordamaktadır.
- e) Hiç yordamamaktadır.

S46. Sizce yaptığınız **konuşma** sınavları genel olarak öğretmen dostu muydu? (Uygulanması ve değerlendirmesi kolay mı?)

- a) Kesinlikle öğretmen dostuydu.
- b) Öğretmen dostuydu.
- c) Ne öğretmen dostu ne değil.
- d) Öğretmen dostu değildi.
- e) Kesinlikle öğretmen dostu değildi.

Appendix-1: (Continue)

S47. Konuşma sınavlarınızın yeterince güvenilir (tutarlı) sonuçlar verdiğini düşünüyor musunuz?

- a) Kesinlikle evet b) Evet, çoğunlukla c) Kararsızım
d) Hayır, pek değil e) Kesinlikle hayır

S48. Yaptığınız dinleme sınavları yeterince geçerli miydi (amaçlananı ölçmüş müydü)?

- a) Kesinlikle evet b) Evet, çoğunlukla c) Kararsızım
d) Hayır, pek değil e) Kesinlikle hayır

S49. Konuşma sınavlarının güvenilirliğini, geçerliliğini veya genel kalitesini artırmak için herhangi bir öneriniz var mı?

- a) Evet b) Hayır

S50. Konuşma sınavlarının güvenilirliğini, geçerliliğini veya genel kalitesini artırmak için önerileriniz nelerdir?

Lütfen kısaca yanıtlayınız: _____.

C: Konuşma Sınavlarının Öğretme ve Öğrenmeye Etkisi

Bu soru grubu, öğretmenlerin sınavların öğrencilere sağladığı katkıyı ve öğretim süreçlerine etkisini nasıl değerlendirdiğini ele alır.

S51. Genel olarak **konuşma** becerisi uygulama sınavından memnuniyetiniz nedir?

- a) Çok memnunum b) Biraz memnunum c) Ne memnunum ne değilim
d) Pek memnun değilim e) Memnun değilim

S52. Konuşma sınav sonuçlarını kullanarak öğrencilere destek sağladınız mı?

- a) Evet b) Hayır

S53. Konuşma sınav sonuçlarını kullanarak öğrencilere ne tür bir destek sağladınız? (**NOT: Birden fazla seçenek işaretleyebilirsiniz?**)

- ☐ Ek ders materyalleri ☐ Bireysel danışmanlık
☐ Grup çalışmaları ☐ Geri bildirim oturumları

S54. Konuşma sınavları, öğrencileri becerilerini geliştirmeleri için ne ölçüde motive etti?

- a) Çok yüksek düzeyde motive etti.
b) Yüksek düzeyde motive etti.
c) Orta düzeyde motive etti.
d) Düşük düzeyde motive etti.
e) Hiç motive etmedi.

S55. Konuşma sınavlarının öğretme ve öğrenmeyi ne kadar etkilediğini düşünüyorsunuz?

- a) Çok olumlu etkiliyor b) Olumlu etkiliyor c) Ne olumlu ne olumsuz
d) Olumsuz etkiliyor e) Çok olumsuz etkiliyor

Appendix-1: (Continue)

S56. Yaptığınız **konusma** sınavlarının öğrencilerin dil gelişimine yeterince katkısı olduğunu düşünüyor musunuz?

- a) Kesinlikle evet b) Evet, çoğunlukla c) Kararsızım
d) Hayır, pek değil e) Kesinlikle hayır

D: Öğretmenlerin Geri Bildirim Kullanımı

Bu soru grubu, öğretmenlerin sınav sonuçlarını nasıl değerlendirdiğini ve öğrencilere geri bildirim sağlama süreçlerini inceler.

S57. Konuşma sınavı değerlendirme sürecinde öğrenci geri bildirimini kullanıyor musunuz?

- a) Evet b) Hayır

S58. Konuşma sınavı değerlendirme sürecinde öğrenci geri bildirimlerini nasıl kullanıyorsunuz? (*NOT: Birden fazla seçenek işaretleyebilirsiniz?*)

- ☐ İyileştirme önerileri sunarak
☐ Öğrencilerin güçlü ve zayıf yönlerini vurgulayarak
☐ İlerleme raporları hazırlayarak
☐ Öğrenci-öğretmen görüşmeleri yaparak
☐ Diğer

S59. Konuşma sınavı değerlendirme sürecinde öğrenci geribildirimini yararlı olacağını düşünüyor musunuz?

- a) Evet b) Hayır

S60. Konuşma sınavı değerlendirme sürecinde öğrenci geri bildirimini nasıl kullanıyorsunuz?

Lütfen kısaca yanıtlayınız: _____.

S61. Öğrencilerin **konusma** becerilerini geliştirmek için onlara herhangi bir geribildirim sağlıyor musunuz?

- a) Evet b) Hayır

S62. Öğrencilerin **konusma** becerilerini geliştirmek için onlara hangi geribildirmiyi sağlıyor musunuz?

- a) Gramer ve telaffuz hatalarını düzeltme
b) İçerik ve anlamı değerlendirme
c) Konuşma akıcılığına ve anlatım becerisine odaklanma
d) Öğrencileri güçlü yönlerine teşvik etme
e) Diğer

Appendix-1: (Continue)

S63. Öğrencilerin **konuşma** becerilerini değerlendirirken hangi kriterleri kullanıyorsunuz? (**NOT:** Birden fazla seçenek işaretleyebilirsiniz?)

- | | |
|--|--|
| ☐ Anlama düzeyi | ☐ Sözcüklerin telaffuzunu anlama |
| ☐ Ana fikri/konuyu kavrama | ☐ Tonlama ve vurgulamayı anlama |
| ☐ Detayları anlama | ☐ Dinleme sonrası çıkarım yapabilme |
| ☐ Kelime dağarcığı zenginliği | ☐ Diğer |
| ☐ Anahtar kavramları/kelimeleri tanım | |

E: Öğretmenlerin Mesleki Gelişim ve Eğitimi Geçmişi

Bu soru grubu, öğretmenlerin konuşma becerilerinin ölçülmesine yönelik aldıkları eğitimi ve mesleki gelişim süreçlerini kapsamaktadır.

S64. Lisans eğitiminizde **konuşma** becerilerinin ölçülmesine yönelik yeterli bir eğitim aldınız mı?

a) Evet

b) Hayır

S65. Lisans eğitiminizde **konuşma** becerilerinin ölçülmesine yönelik aldığınız bu eğitimi nasıl değerlendiriyorsunuz?

a) Çok iyi

b) İyi

c) Orta

d) Yetersiz

e) Oldukça yetersiz

S66. Lisans sonrasında **konuşma** becerilerinin ölçülmesine yönelik herhangi bir mesleki etkinliğe (konferans, seminer, webinar vb.) katıldınız?

a) Evet

b) Hayır

S67. **Konuşma** becerilerinin ölçülmesine yönelik bir eğitime katılmak ister miydiniz? Böylesi bir etkinliği ne derece yararlı olacağını düşünüyorsunuz?

a) Evet, katılmak isterim ve çok yararlı olacağını düşünüyorum

b) Evet, katılmak isterim ve yararlı olacağını düşünüyorum.

c) Kararsızım.

d) Hayır, katılmak istemem ama yararlı olacağını düşünüyorum.

e) Hayır, katılmak istemem ve yararlı olacağını düşünmüyorum.

S68. Şu an **konuşma** becerilerinin ölçülmesine yönelik aşağıdaki mesleki etkinliklerinden hangisi sizin için daha verimli olur? (**NOT:** Birden fazla seçenek işaretleyebilirsiniz?)

☐ Konferanslar

☐ Seminerler

☐ Webinarlar

☐ Panel Tartışmaları

☐ Sunumlar

☐ Münazaralar

☐ Eğitim Yarışmaları

☐ Çevrimiçi Dersler

Appendix 2: Semi-Structured Interview Questions

1. Milli Eğitim Bakanlığı'nın İngilizce sınavlarında yazılı sınavların yanı sıra **dinleme ve konuşma** sınavlarını eklemesi hakkında ne düşünüyorsunuz?
2. Bu yeniliği genel olarak nasıl karşılıyorsunuz? Öğrencilerin değerlendirilmesi bakımından daha kapsamlı bir yaklaşım olduğunu düşünüyor musunuz?
3. **Dinleme ve konuşma** sınavlarının öğrenciler için avantajları neler olabilir?
4. **Dinleme ve konuşma** sınavlarının öğrenci başarısı üzerinde nasıl bir etkisi olabilir? Yazılı sınavlarla karşılaştırıldığında farklar nelerdir?
5. **Dinleme ve konuşma** sınavlarında karşılaştığınız zorluklar nelerdir? Öğrenciler veya öğretmenler açısından herhangi bir zorluk yaşıyor mu?
6. **Dinleme ve konuşma** sınavlarının öğrencilerin genel dil becerilerini değerlendirme açısından güçlü yönleri ve zayıf yönleri nelerdir?
7. Sizce **dinleme ve konuşma** becerilerinden hangisinin ölçülmesi öğrencinin dil gelişimine daha büyük katkı sağladığını düşünüyorsunuz?
8. Bu tür sınavlar, öğrencilerin İngilizceye daha etkili bir şekilde öğrenmelerini teşvik edebilir mi? Ne gibi faydalar sağlayabilir?
9. **Dinleme ve konuşma** sınavlarının öğretmenlerin işini kolaylaştırdığı veya zorlaştırdığı noktalar nedir?
10. **Dinleme ve konuşma** sınavlarının öğrencilerin özgüvenlerini ve dil kullanımını artırma konusunda etkili olup olmadığına dair gözlemlerinizi nelerdir?
11. Gelecekte bu sınav türlerinin daha da geliştirilmesi adına önerileriniz var mı?

Appendix 3: Research Ethics Committee Approval



HİZMETE ÖZEL

T.C.

KARADENİZ TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Hukuk Müşavirliği

18.07.2024 17:48
E-82554930-050.01.04-537079
04758768



Sayı : E-82554930-050.01.04-537079 /401
Konu : Rabia ACAR'ın Etik Kurul Kararı

18.07.2024

Sn. Rabia ACAR

İlgi : 18.07.2024 tarihli ve E-26014373-050.01.04-536442 sayılı yazı.

KTÜ-Sosyal Bilimler Enstitüsü Batı Dilleri ve Edebiyatı Anabilim Dalı Uygulamalı Dil Bilimi Tezli Yüksek Lisans Programı öğrencisi Rabia ACAR'ın "**Voices from EFL Teachers Integrating Listening and Speaking Components in Assessment (İngilizce Öğretmenlerinin Dinleme ve Konuşma Becerilerinin Değerlendirmeye Dahil Edilmesine İlişkin Düşünceleri**" başlıklı yüksek lisans tezi kapsamında çalışma için gerekli olan Etik Kurul incelemesi yapılmış ve başvurunuza onay verilmiştir.

Bilgilerinize ve gereğini rica ederim.

Prof. Dr. Akif CİNEL
Rektör Yardımcısı

Doğrulama Kod: BC0C9EAC-09FB-49E7-B1C1-81870C6CF84E

Bu belge güvenli elektronik imza ile onaylanmıştır.

HİZMETE ÖZEL

Doğrulama Adres: <https://www.turkiye.gov.tr/karadeniz-teknik-universitesi-cbys>

61080 – Trabzon /
TÜRKİYE
Tel: 04623772107

Faks: 04623774399

www.ktu.edu.tr

Ayrıntılı Bilgi İçin İrtibat
Belgin ŞEN
hukukmusavirligi@ktu.edu.tr

Sayfa
1 / 1



Appendix-3: Continue



HİZMETE ÖZEL

T.C.

KARADENİZ TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Sosyal ve Beşeri Bilimler Etik Kurulu

18.07.2024 09:30
E-26014373-050.01.04-536442
04756510



Sayı : E-26014373-050.01.04-536442
Konu : Rabia ACAR'ın Etik Kurul Kararı

18.07.2024

REKTÖRLÜK MAKAMINA
Hukuk Müşavirliği

İlgi : 26.06.2024 tarihli ve E-90783813-199-15056 sayılı yazı.

KTÜ – Sosyal Bilimler Enstitüsü Batı Dilleri ve Edebiyatı Anabilim Dalı öğretim üyesi Dr. Öğr. Üyesi Hasan SAĞLAMEL'in akademik danışmanlığında Uygulamalı Dil Bilimi Tezli Yüksek Lisans Programı öğrencisi 382659 numaralı Rabia ACAR'ın yürüttüğü "Voices from EFL Teachers Integrating Listening and Speaking Components in Assessment (İngilizce Öğretmenlerinin Dinleme ve Konuşma Becerilerinin Değerlendirmeye Dahil Edilmesine İlişkin Düşünceleri" başlıklı yüksek lisans tezi kapsamında "Etik Kurul Belgesi" ile ilgili olarak kurulumuza yaptığı başvuru değerlendirilerek uygun bulunmuş ve sonucu Ek'te gönderilmiştir.

İlgili öğrencinin bilgilendirilmesi hususunda gereğini arz ederim.

Prof. Dr. Abdülkadir PEHLİVAN
Başkan

Ek: Toplantı Tutanağı (3 Sayfa)



Çetin Kaya GÜMRÜK
Hukuk Müşavirliği

Doğrulama Kod: 7894F4CD-8367-47EC-B8A1-6FB117DBA7B5

Bu belge güvenli elektronik imza ile onaylanmıştır.

HİZMETE ÖZEL

Doğrulama Adres: <https://www.turkiye.gov.tr/karadeniz-teknik-universitesi-cbya>

61080 – Trabzon /
TÜRKİYE
Tel: 0 (462) 377 43 36

Faks: 0 (462) 325 32 05

www.ktu.edu.tr

Ayrıntılı Bilgi İçin İrtibat
Nurcan SAHİN
genelsekreterlik@ktu.edu.tr

Sayfa
1 / 1



Appendix-4: Research Permission License from the Ortahisar Directorate of National Education



TRABZON ORTAHİSAR İlçe Milli Eğitim Müdürlüğüne



Başvuru No: MEB.TT.2024.000588

Uygulama Yapılacak MEB Teşkilatının Kurum Kodu: 753131

Adı Soyadı: RABİA ACAR

Araştırmanın Adı: İngilizce Öğretmenlerinin Dinleme ve Konuşma Becerilerinin Değerlendirmeye Dahil Edilmesine İlişkin Düşünceleri

Araştırmanın Niteliği: Yüksek Lisans Tezi

Araştırmanın Örneklem / Çalışma Grubu: Öğretmen

Uygulama Yapılacak MEB Teşkilatı: TRABZON ORTAHİSAR İlçe Milli Eğitim Müdürlüğü

Uygulama Yapılacak Birim: İlçe Milli Eğitim Müdürlüğü

Uygulama Yapılacak İl: TRABZON

Veri Toplama Aracının Başlığı: İNGİLİZCE ÖĞRETMENLERİNİN DİNLEME VE KONUŞMA BECERİLERİNİN DEĞERLENDİRMEYE DAHİL EDİLMESİNE İLİŞKİN DÜŞÜNCELERİ

Araştırma Uygulama İzininin Kabul Tarihi: 05.09.2024

Araştırmanın Uygulama İzininin Bitiş Tarihi: 05.09.2025

Yukarıda kimliği yazılı araştırmacı "Araştırma Uygulama İzinleri Genelgesine (2024/41)" göre belirtilen kapsamda araştırmasını yapmayı taahhüt etmiştir. Araştırmacının bilgi ve belgelerinin uygunluğu kontrol edilmiş olup araştırma uygulama izni TRABZON İl Milli Eğitim Müdürlüğü tarafından onaylanmıştır.

NOT: Okul/kurum yöneticileri tarafından "Araştırma Uygulama İzni" belgesinin ve veri toplama araçlarının (araçlardaki maddelerinin) modülde yer alan belge ve araçlarla aynı olduğu kontrol edilmelidir. Aynı olmadığı durumda araştırma uygulama izni verilmeyecektir.

* Başvuru detayını görüntülemek ve belgeyi doğrulamak için 'https://arastirmaizinleri.meb.gov.tr/belge-dogrula' bağlantısını kullanınız.

- Araştırma Uygulama İzinleri Başvuru ve Değerlendirme Modülü -

CIRRICULUM VITAE

She completed her high school education at Fatsa Anatolian High School in 2008. In 2013, she graduated from the Department of English Language and Literature at Karadeniz Technical University (KTU). During her undergraduate studies, she participated in the Erasmus Student Exchange Program and studied at the Department of Foreign Languages and Literature at the University of Salerno in Italy. Following her graduation, she briefly worked as an English instructor at TEOL Language Schools and then as a contract English teacher at Kaşüstü Multi-Program Anatolian High School. In the 2013-2014 academic year, she completed the Pedagogical Formation Program at the Faculty of Education at KTU. In 2015, she was appointed as an English teacher by the Ministry of National Education of the Republic of Türkiye. She first served at Andaçlı Primary School in Patnos, Ağrı (2015-2016) and later at Cumhuriyet Middle School in Ortahisar, Trabzon (2016-2017). Since 2017, she has been continuing her teaching career at Mimar Sinan Middle School in Ortahisar, Trabzon. ACAR is currently pursuing a master's degree in Applied Linguistics at the Department of Western Languages and Literature at Karadeniz Technical University.

ACAR is married and the mother of a son.